

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 14:09:48 / Search time 34 Seconds
(without alignments)
478.156 Million cell updates/sec

Title: US-09-822-540A-1

Perfect score: 346
Sequence: 1 ADRAAVPIVNLKDELLFPSW.....DVLRHPTWPKSVWHS DPN 63

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

SPTRMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	346	100.0	187	4	OBWX15
2	346	100.0	261	4	OBNG19
3	346	100.0	816	4	OBNG19
4	298	86.1	102	4	OBNG19
5	294	85.0	160	11	O9CMT2
6	294	85.0	184	11	O9JRK63
7	294	85.0	1140	11	O61434
8	294	85.0	1774	11	O62001
9	284	82.1	171	11	O9WMT5
10	284	82.1	226	11	O9QZD2
11	255	72.3	1344	13	O93419
12	250	72.3	1315	13	O8QHL9
13	233	67.3	1307	13	O8UPF7
14	219	63.9	1361	13	OBAMC6
15	197	56.9	1367	11	O35206
16	196	56.6	1367	11	O9EOD9

17	193	55.8	1388	4	O9Y4W4	O9Y4W4 homo sapien
18	169.5	49.0	299	5	OBMS33	OBMS33 drosophila
19	169.5	49.0	792	5	OBMT89	OBMT89 drosophila
20	140	40.5	650	5	O17866	O17866 caenorhabdi
21	140	40.5	778	5	O9U9K5	O9U9K5 caenorhabdi
22	140	40.5	1117	5	O9U9K7	O9U9K7 caenorhabdi
23	140	40.5	1154	5	O810G6	O810G6 caenorhabdi
24	70.5	20.4	317	4	O43467	O43467 homo sapien
25	67.5	19.5	614	4	O03989	O03989 homo sapien
26	65	18.8	958	5	O8SX15	O8SX15 drosophila
27	64	18.5	644	17	O9YAI3	O9YAI3 aeropyrum
28	63.5	18.4	266	10	O8M159	O8M159 brassica oi
29	63.5	18.4	847	16	O8Y0C4	O8Y0C4 ralbionia s
30	63	18.2	336	2	O9ZG16	O9ZG16 streptomyce
31	62.5	18.1	197	17	O8ZTQ1	O8ZTQ1 pyrobaculum
32	62.5	18.1	467	5	O9XX44	O9XX44 caenorhabdi
33	62.5	18.1	620	16	O11526	O11526 bacillus su
34	62	17.9	102	4	O8WY95	O8WY95 homo sapien
35	62	17.9	126	17	O8TU72	O8TU72 methanosarc
36	62	17.9	325	4	O9Y247	O9Y247 homo sapien
37	62	17.9	419	5	P91096	P91096 caenorhabdi
38	61.5	17.8	362	7	O9MXH0	O9MXH0 pan troglod
39	61.5	17.8	1172	12	O9E1Y8	O9E1Y8 cercopithec
40	61	17.6	738	11	O9Z213	O9Z213 ratius norv
41	60.5	17.5	236	3	O8X1J5	O8X1J5 cryptococcu
42	60.5	17.5	862	10	O9FT17	O9FT17 lycopersico
43	60	17.3	258	17	P94920	P94920 methanosarc
44	60	17.3	258	17	O9C4S2	O9C4S2 methanosarc
45	60	17.3	260	17	O8Q0C2	O8Q0C2 methanosarc

ALIGNMENTS

RESULT 1

OBWX15 PRELIMINARY; PRT: 187 AA.

AC OBWX15; 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Collagen XVIII (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_Taxid=9606;

RM [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21409408; PubMed=11517600;

RA Feng Y., Cui L.B., Liu C.X., Ma Q.J.

RT "Inhibition effect in vitro of purified endostatin expressed in Pichia pastoris".

RL Sheng Wu Gong Cheng Xue Bao 17:278-282(2001).

DR EMBL; AF416592; AAL37720.1; ..

FT NON_TER 1

SO SEQUENCE 187 AA; 20448 MW; 72B1047D85838CD3 CRC64;

Query Match 100.0%; Score 346; DB 4; Length 187;

Best Local Similarity 100.0%; Pred. No. 3e-34; 0; Indels 0; Gaps 0;

Matches 63; Conservative 0; Mismatches 0;

QY 1 ADRAAVPIVNLKDELLFPSWEGPLKRGARIFSPDGKDVLRHPTWPKSVWHS 60

DB 68 ADRAAVPIVNLKDELLFPSWEGPLKRGARIFSPDGKDVLRHPTWPKSVWHS 127

QY 61 DPN 63

DB 128 DPN 130

RESULT 2

OBNG19 PRELIMINARY; PRT: 261 AA.

```

AC Q8NG19;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DE Endostatin variant (Fragment).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Submitted (JUN-2000) to the EMBL/GenBank/DDJ databases.
RL EMBL; AF282883; AAM52249.1; -.
SQ SEQUENCE 261 AA; 26745 MW; CA60C920AF3E90E5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 346; DB 4; Length 261;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPGSEGLPKRGARIFSGDKDVLRHPTWPKSVWHS 60
DB 142 ADRAAVPIVNLKDELLFPGSEGLPKRGARIFSGDKDVLRHPTWPKSVWHS 201
QY 61 DPN 63
DB 202 DPN 204

RESULT 3
Q8N4S4
ID Q8N4S4; PRELIMINARY; PRT; 816 AA.
AC Q8N4S4;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DE Endostatin variant (Fragment).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Submitted (JUL-2002) to the EMBL/GenBank/DDJ databases.
RL EMBL; BC033715; AAH33715.1; -.
DE InterPro: IPR000087; Collagen.
DR Pfam: PF01391; Collagen; 6.
DR ProDom: PD000007; Collagen; 1.
KW Collagen.
FT NON_TER
SQ SEQUENCE 816 AA; 82553 MW; 5D539B2946694F86 CRC64;

Query Match
Best Local Similarity 100.0%; Score 346; DB 4; Length 816;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPGSEGLPKRGARIFSGDKDVLRHPTWPKSVWHS 60
DB 697 ADRAAVPIVNLKDELLFPGSEGLPKRGARIFSGDKDVLRHPTWPKSVWHS 756
QY 61 DPN 63
DB 757 DPN 759

RESULT 4
Q86T70
ID Q86T70; PRELIMINARY; PRT; 102 AA.
AC Q86T70;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DE Endostatin variant (Fragment).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Submitted (JAN-2001) to the EMBL/GenBank/DDJ databases.
RL EMBL; AF333247; AAK50626.1; -.
FT NON_TER
SQ SEQUENCE 102 AA; 11147 MW; ECAC47AA6420947D CRC64;

Query Match
Best Local Similarity 86.1%; Score 298; DB 4; Length 102;
Matches 55; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 IYNLDELLFPGSEGLPKRGARIFSGDKDVLRHPTWPKSVWHS 63
DB 1 IYNLDELLFPGSEGLPKRGARIFSGDKDVLRHPTWPKSVWHS 56
QY 61 DPN 63
DB 202 DPN 204

RESULT 5
Q8CRT2
ID Q8CRT2; PRELIMINARY; PRT; 160 AA.
AC Q8CRT2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DE Endostatin variant (Fragment).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Submitted (JUN-2001) to the EMBL/GenBank/DDJ databases.
RL EMBL; AF333247; AAK50626.1; -.
FT NON_TER
SQ SEQUENCE 160 AA; 17725 MW; 60F853D777C375D2 CRC64;

Query Match
Best Local Similarity 85.0%; Score 294; DB 11; Length 160;
Matches 51; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

```

```

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Endostatin variant (Fragment).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Deininger M.H., Trautmann K., Schluessener H.J.;
RT "Endostatin promotes delayed secondary damage following traumatic
brain injury."
RL Submitted (JAN-2001) to the EMBL/GenBank/DDJ databases.
DR EMBL; AF333247; AAK50626.1; -.
FT NON_TER
SQ SEQUENCE 102 AA; 11147 MW; ECAC47AA6420947D CRC64;

Query Match
Best Local Similarity 98.2%; Score 298; DB 4; Length 102;
Matches 55; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 IYNLDELLFPGSEGLPKRGARIFSGDKDVLRHPTWPKSVWHS 63
DB 1 IYNLDELLFPGSEGLPKRGARIFSGDKDVLRHPTWPKSVWHS 56
QY 61 DPN 63
DB 202 DPN 204

RESULT 5
Q8CRT2
ID Q8CRT2; PRELIMINARY; PRT; 160 AA.
AC Q8CRT2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DE Endostatin variant (Fragment).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Submitted (JUN-2001) to the EMBL/GenBank/DDJ databases.
RL EMBL; AF333247; AAK50626.1; -.
FT NON_TER
SQ SEQUENCE 160 AA; 17725 MW; 60F853D777C375D2 CRC64;

Query Match
Best Local Similarity 85.0%; Score 294; DB 11; Length 160;
Matches 51; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      1 ADRAVPIVNLKDELFPSEWALFSGSEGLPKGARIFSPDGDVLRHPTWPKSVWHS 60
DB      40 ADRGVPIVNLKDEVLPSEWDLFSGSQGLQPGARIFSPDGDVLRHPTWPKSVWHS 99
QY      61 DPN 63
DB      100 DPS 102

RESULT 6
QYK63   PRELIMINARY; PRT; 184 AA.
AC      09UK63:
DT      01-OCT-2000 (TREMBLrel. 15, Created)
DT      01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE      Endostatin (Fragment).
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-Chinese Kunming;
RA      Jia S., Zhu F., King G., Yu Y., Duan C., Xiu R.-J., He F.;
RT      "Anticancer treatment of targeted fusion protein delivery to tumor
RT      neovasculature.";
RL      Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF257775; AAF69009.1; -.
FT      NON_TER
FT      NON_TER
SQ      SEQUENCE 184 AA; 20376 MW; AC06F9D8D103412A CRC64;

Query Match
Best Local Similarity 85.0%; Score 294; DB 11; Length 184;
Matches 51; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY      1 ADRAVPIVNLKDELFPSEWALFSGSEGLPKGARIFSPDGDVLRHPTWPKSVWHS 60
DB      64 ADRGVPIVNLKDEVLPSEWDLFSGSQGLQPGARIFSPDGDVLRHPTWPKSVWHS 123
QY      61 DPN 63
DB      124 DPS 126

RESULT 7
QYK63   PRELIMINARY; PRT; 1140 AA.
AC      061434:
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE      01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE      Collagen (Fragment).
GN      COL15A1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      Abe N., Muragaki Y., Yoshioke H., Inoue H., Ninomiya Y.;
RT      "Identification of a novel collagen chain represented by extensive
RT      interruptions in the triple-helical region.";
RL      Cell. Mol. Biol. Res. 196; 576-582 (1993).
DR      EMBL; D17546; BAA04483.1; -.
DR      HSSP; P39061; 1KOE.
DR      MGD; MGI:88449; Col15a1.
DR      InterPro; IPR000087; Collagen.
DR      InterPro; IPR003129; TSPN.
DR      Pfam; PF01391; Collagen; 8.
DR      Pfam; PF02210; TSPN; 1.

```

```

DR      Prodom; PD000007; Collagen; 1.
KW      Collagen.
FT      NON_TER
SQ      SEQUENCE 1140 AA; 115156 MW; 8B0C7E6862B3BDFE CRC64;

Query Match
Best Local Similarity 85.0%; Score 294; DB 11; Length 1140;
Matches 51; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY      1 ADRAVPIVNLKDELFPSEWALFSGSEGLPKGARIFSPDGDVLRHPTWPKSVWHS 60
DB      1020 ADRGVPIVNLKDEVLPSEWDLFSGSQGLQPGARIFSPDGDVLRHPTWPKSVWHS 1079
QY      61 DPN 63
DB      1080 DPS 1082

RESULT 8
QYK63   PRELIMINARY; PRT; 1774 AA.
AC      062001:
ID      062001; Q60672;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      Procollagen, type XVIII, alpha 1 precursor (XVIII) collagen
DE      (Procollagen, type XVIII, alpha 1) (Alpha-1 type XVIII collagen).
GN      COL18A1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE-PRIMARY TAIL CULTURE;
RX      MEDLINE=94245707; PubMed=8186673;
RA      Rehn M., Hunklika E., Phlajaniemi T.;
RT      "Primary structure of the alpha 1 chain of mouse type XVIII collagen,
RT      partial structure of the corresponding gene, and comparison of the
RT      alpha 1(XVIII) chain with its homologue, the alpha 1(XV) collagen
RT      chain.";
RL      J. Biol. Chem. 269; 13929-13935 (1994).
RN      [2]
RP      SEQUENCE OF 1-562 FROM N.A.
RX      MEDLINE=94240112; PubMed=8183894;
RA      Rehn M., Phlajaniemi T.;
RT      "Alpha 1(XVIII), a collagen chain with frequent interruptions in the
RT      collagenous sequence, a distinct tissue distribution, and homology
RT      with type XV collagen.";
RL      Proc. Natl. Acad. Sci. U.S.A. 91; 4234-4238 (1994).
RN      [3]
RP      SEQUENCE OF 1-562 FROM N.A.
RX      MEDLINE=95181468; PubMed=7876242;
RA      Rehn M., Phlajaniemi T.;
RT      "Identification of three N-terminal ends of type XVIII collagen chains
RT      and tissue-specific differences in the expression of the corresponding
RT      transcripts. The longest form contains a novel motif homologous to rat
RT      and Drosophila frizzled proteins.";
RL      J. Biol. Chem. 270; 4705-4711 (1995).
DR      EMBL; U03715; AAC52903.1; -.
DR      EMBL; U03716; AAC52903.1; JOINED.
DR      EMBL; U03718; AAC52903.1; JOINED.
DR      EMBL; U34607; AAC52903.1; JOINED.
DR      EMBL; U34608; AAC52903.1; JOINED.
DR      EMBL; U34609; AAC52903.1; JOINED.
DR      EMBL; U34610; AAC52903.1; JOINED.
DR      EMBL; U34611; AAC52903.1; JOINED.
DR      EMBL; U34612; AAC52903.1; JOINED.
DR      EMBL; U34613; AAC52903.1; JOINED.
DR      EMBL; U11637; AAC52179.1; -.
DR      HSSP; P39061; 1KOE.
DR      MGD; MGI:88451; Col18a1.
DR      InterPro; IPR000087; Collagen.

```

DR InterPro: IPR000024; Fz domain.
 DR InterPro: IPR001791; Laminin_G.
 DR InterPro: IPR003129; TSPN.
 DR Pfam: PF01391; Collagen; 8.
 DR Pfam: PF01392; Fz; 1.
 DR Pfam: PF02210; TSPN; 1.
 DR ProDom: PD000007; Collagen; 1.
 DR SMART: SM00063; FRI; 1.
 DR SMART: SM00282; LamG; 1.
 DR SMART: SM00210; TSPN; 1.
 DR PROSITE: PS50038; FZ; 1.
 KW Collagen; Signal.
 SQ SEQUENCE 1774 AA; 182229 MW; CP49B9C9E88EF232 CRC64;

Query Match 85.0%; Score 294; DB 11; Length 1774;
 Best Local Similarity 81.0%; Pred. No. 8.3e-27;
 Matches 51; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELFPSEWALFSGSEGLPKGARIFFPDGKDVLRHPTWPKSVWHS 60
 DB 1654 ADRAAVPIVNLKDELFPSEWALFSGSQGLGARIFFPDGKDVLRHPTWPKSVWHS 1713
 QY 61 DPN 63
 DB 1714 DPS 1716

RESULT 9

Q9WDM5 PRELIMINARY; PRT; 171 AA.
 AC Q9WDM5; 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Collagen type XVIII, alpha (1) chain (fragment).
 GN COL18A1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Jia J.D., Bauer M., Eberpaecher U., Donner P., Schuppan D.;
 RT "Partial 3'-terminal cDNA sequence of rat collagen XVIII/endostatin.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Jia J.D., Bauer M., Sedlacek N., Ruenli M., Riecken E.O., Schuppan D.;
 RT "Temporospatial expression of collagen XVIII/endostatin in acute and
 chronic liver injuries";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ236873; CAB44263.1; -.
 DR HSSP: P39061; IKOE.
 FT NON_TER 1
 FT NON_TER 171
 SQ SEQUENCE 171 AA; 18933 MW; 81BE2EE3FC2C8E72 CRC64;

Query Match 82.1%; Score 284; DB 11; Length 171;
 Best Local Similarity 79.4%; Pred. No. 9.8e-27;
 Matches 50; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELFPSEWALFSGSEGLPKGARIFFPDGKDVLRHPTWPKSVWHS 60
 DB 56 ADRAAVPIVNLKDELFPSEWALFSGSQGLGARIFFPDGKDVLRHPTWPKSVWHS 115
 QY 61 DPN 63
 DB 116 DPS 118

RESULT 10

-

Q9QZD2 PRELIMINARY; PRT; 226 AA.
 AC Q9QZD2; 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Collagen XVIII (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RC MEDLINE=20227226; PubMed=10766159;
 RA Perletti G., Concarri P., Giardini R., Marra E., Piccinini F.,
 RA Folkmann J., Chen L.;
 RT "Antitumor activity of endostatin against carcinogen-induced rat
 primary mammary tumors";
 RL Cancer Res. 60:1793-1796(2000).
 DR EMBL: AF189709; AAFO0975.1; -.
 DR HSSP: P39061; IKOE.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 226 AA; 25350 MW; 38B83C0486C0B949 CRC64;

Query Match 82.1%; Score 284; DB 11; Length 226;
 Best Local Similarity 79.4%; Pred. No. 1.3e-26;
 Matches 50; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELFPSEWALFSGSEGLPKGARIFFPDGKDVLRHPTWPKSVWHS 60
 DB 106 ADRAAVPIVNLKDELFPSEWALFSGSQGLGARIFFPDGKDVLRHPTWPKSVWHS 165
 QY 61 DPN 63
 DB 166 DPS 168

RESULT 11

O9J419 PRELIMINARY; PRT; 1344 AA.
 AC O9J419; 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Collagen XVIII precursor.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98411346; PubMed=9738008;
 RA Halter W., Dong S., Schurer B., Cole G.J.;
 RT "Collagen XVIII is a basement membrane heparan sulfate proteoglycan";
 RL J. Biol. Chem. 273:25404-25412(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Halter W., Dong S.;
 RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF083440; AAC33294.2; -.
 DR HSSP: P39061; IKOE.

DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR001791; Laminin_G.
 DR InterPro: IPR003129; TSPN.
 DR Pfam: PF01391; Collagen; 8.
 DR Pfam: PF02210; TSPN; 1.
 DR ProDom: PD000007; Collagen; 1.
 DR SMART: SM00282; LamG; 1.
 DR SMART: SM00210; TSPN; 1.
 KW Collagen; Signal.
 FT SIGNAL 1
 FT SIGNAL 26
 * POTENTIAL.

SQ SEQUENCE 1344 AA; 137402 MW; 7AA366E4FE940CCD CRC64;
 Query Match 73.7%; Score 255; DB 13; Length 1344;
 Best Local Similarity 72.1%; Pred. No. 3.5e-22;
 Matches 44; Conservative 9; Mismatches 8; Indels 0; Gaps 0;
 QY 1 ADRAAPVPIVNLKDELLFPSWEALFSGSEGLKPGARIFSGDKDVL RHPTWPKSWHGS 60
 DB 1224 ADRTAVPIVNLDEVLFSNWELFTGSEAPLRAGARILSPDGRDIIQDSAMWPKSWHGS 1283
 QY 61 D 61
 DB 1284 D 1284
 RESULT 12
 Q8QHL9 PRELIMINARY; PRT; 1315 AA.
 ID Q8QHL9;
 AC Q8QHL9;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Type XVIII collagen alpha1 chain.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 [1]
 RP SEQUENCE FROM N.A.
 RA Ichino T., Sekimizu K., Natori S., Kubo T.;
 RT "Identification and characterization of genes expressed selectively in
 the regenerating tail of *Xenopus laevis* tadpole."
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB047066; BAB84674.1; -
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR003129; TSPN.
 DR Pfam; PF01391; Collagen; 7.
 DR Pfam; PF02210; TSPN; 1.
 DR ProDom; PD000007; Collagen; 1.
 DR SMART; SM00210; TSPN; 1.
 KW Collagen.
 SQ SEQUENCE 1315 AA; 134946 MW; 0C56C235DE058365 CRC64;
 Query Match 72.3%; Score 250; DB 13; Length 1315;
 Best Local Similarity 72.1%; Pred. No. 1.4e-21;
 Matches 44; Conservative 8; Mismatches 9; Indels 0; Gaps 0;
 QY 1 ADRAAPVPIVNLKDELLFPSWEALFSGSEGLKPGARIFSGDKDVL RHPTWPKSWHGS 60
 DB 1195 ADROSVOIVNLDEVLVDNWESLFGSGEAKMRPGARIFSGDKDVLDTPTWPKSWHGS 1254
 QY 61 D 61
 DB 1255 D 1255
 RESULT 13
 Q8JFF7 PRELIMINARY; PRT; 1307 AA.
 ID Q8JFF7;
 AC Q8JFF7;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Type XVIII collagen short variant.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=22166979; PubMed=12175494;

RA Elamaa H., Peterson J., Pihlajaniemi T., Deestree O.;
 RT "Cloning of three variants of type XVIII collagen and their expression
 patterns during *Xenopus laevis* development."
 RL Mech. Dev. 114:109-113(2002).
 DR EMBL; AY052763; AAL14257.1; -
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR003129; TSPN.
 DR Pfam; PF01391; Collagen; 6.
 DR Pfam; PF02210; TSPN; 1.
 DR SMART; SM00210; TSPN; 1.
 KW Collagen.
 SQ SEQUENCE 1307 AA; 134153 MW; D53EDBF83DE34976 CRC64;
 Query Match 67.3%; Score 233; DB 13; Length 1307;
 Best Local Similarity 67.2%; Pred. No. 1.6e-19;
 Matches 41; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
 QY 1 ADRAAPVPIVNLKDELLFPSWEALFSGSEGLKPGARIFSGDKDVL RHPTWPKSWHGS 60
 DB 1187 ADROSVOIVNLDEVLVDNWESLFGSGEAKMRPGARILSPDGRDVLDTPTWPKSWHGS 1246
 QY 61 D 61
 DB 1247 D 1247
 RESULT 14
 Q8AMC6 PRELIMINARY; PRT; 361 AA.
 ID Q8AMC6;
 AC Q8AMC6;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Collagen XVIII (Fragment).
 GN COL18A1.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_TaxID=7955;
 [1]
 RP SEQUENCE FROM N.A.
 RA Hatlex Z., Morvan-Dubois G., Thisse B., Garrone R., Le Guelliec D.;
 RT "Sequence and embryonic expression of collagen XVIII NC11 domain
 (endostatin) in the zebrafish."
 RL Submitted (Jul-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ494837; CAD38825.1; -
 FT NON_TER
 SQ SEQUENCE 361 AA; 40222 MW; 3C5A0F8479D26735 CRC64;
 Query Match 63.3%; Score 219; DB 13; Length 361;
 Best Local Similarity 62.3%; Pred. No. 1.9e-18;
 Matches 38; Conservative 10; Mismatches 13; Indels 0; Gaps 0;
 QY 1 ADRAAPVPIVNLKDELLFPSWEALFSGSEGLKPGARIFSGDKDVL RHPTWPKSWHGS 60
 DB 241 SDRFTLPIVNLKDVLFSSWESLFGSGESRMKDNAPVPSFGDRDVLRSAMPEKSWHGS 300
 QY 61 D 61
 DB 301 D 301
 RESULT 15
 Q35206 PRELIMINARY; PRT; 1367 AA.
 ID Q35206;
 AC Q35206;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Type XV collagen.
 GN COL15A1.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97480713; PubMed=9339358;
 RA Hagg P.M., Horelli-Kuitunen N., Eklund L., Palotie A.,
 RA Pihlajaniemi T.;
 RT "Cloning of mouse type XV collagen sequences and mapping of the
 RT corresponding gene to 4B1-3. Comparison of mouse and human alpha 1
 RT (XV) collagen sequences indicates divergence in the number of small
 RT collagenous domains.";
 RL Genomics 45:31-41(1997).
 DR EMBL; AF011450; AAC53387.1; -.
 DR HSSP; P39061; IKOE.
 DR MGD; MG1:88449; COL15a1.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001791; Laminin_G.
 DR InterPro; IPR003129; TSPN.
 DR Pfam; PF01391; Collagen; 5.
 DR Pfam; PF02210; TSPN; 1.
 DR ProDom; PD000007; Collagen; 1.
 DR SMART; SM00282; LamG; 1.
 DR SMART; SM00210; TSPN; 1.
 KM Collagen.
 ZO SEQUENCE 1367 AA; 140525 MW; A483A1254AF3AEEC CRC64;

Query Match 56.9%; Score 197; DB 11; Length 1367;
 Best Local Similarity 52.4%; Pred. No. 4.2e-15;
 * Matches 33; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELFPSEALFSGSEGPLKPGARIFSPDGKDLRHPTWPKSVYHGS 60
 DB 1247 AERFGLPYVNLKQVLFNNWDSIFSGDGGQFNTHIPYSPGRDVMTPSWPQKVYVHGS 1306
 QY 61 DPN 63
 DB 1307 NPH 1309

Search completed: September 10, 2003, 14:11:14
 Job time : 37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 14:09:48 ; Search time 11 Seconds
(without alignments)
269.335 Million cell updates/sec

Title: US-09-822-540A-1

Perfect score: 346
Sequence: 1 ADRAAVPIVNLKDELLPSPW.....DVLRRHPTWPKSVHGSDPN 63

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt 41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	346	100.0	1516	1 CAIH_HUMAN	P39060 homo sapien
2	294	85.0	1527	1 CAIH_MOUSE	P39061 mus musculu
3	193	55.8	1388	1 CAIE_HUMAN	P39059 homo sapien
4	63.5	18.4	266	1 CB21_SINAI	P13851 sinapis alb
5	61	17.6	683	1 AC51_KLULIA	O60011 kluyveromyc
6	61	17.6	1045	1 PRTS_SBRMA	P09499 serratia ma
7	60	17.3	694	1 TKT1_RHIME	P58333 rhizobium m
8	58.5	16.9	1184	1 ALAC_ARATH	P57792 arabidopsis
9	58.5	16.9	1720	1 FTSH_CHLVU	P56369 chlorella v
10	58	16.8	662	1 AC5A_SCHPO	P78773 schizosacch
11	57.5	16.6	553	1 FXCI_HUMAN	O12948 homo sapien
12	57	16.5	683	1 KPCL_MOUSE	P23298 mus musculu
13	56.5	16.3	770	1 TRPG_ASPNG	P05328 aspergillus
14	56.5	16.3	884	1 YP67_MYCTU	O50654 mycobacteri
15	56	16.2	651	1 NYM_METUA	O58659 methanococc
16	55.5	16.0	727	1 KCRF_HUMAN	P28331 homo sapien
17	55.5	16.0	1174	1 KCRF_STRPU	O14320 strongyloce
18	55	15.9	339	1 XAP5_HUMAN	O14320 homo sapien
19	55	15.9	380	1 OPS4_DROES	P29404 drosophila
20	55	15.9	440	1 YA33_SCHPO	O09710 schizosacch
21	55	15.9	444	1 PAPA_CANFA	O28262 c plateliet
22	55	15.9	545	1 AIRE_HUMAN	O43918 homo sapien
23	55	15.9	673	1 XRP2_PIG	O95333 sus scrofa
24	55	15.9	683	1 KPCL_RAT	O64617 rattus norv
25	54.5	15.8	1861	1 MAP2_RAT	P15146 rattus norv
26	54	15.6	215	1 COAT_PMY	P16596 papaya mosa
27	54	15.6	434	1 KP58_MOUSE	P24788 mus musculu
28	54	15.6	436	1 KP58_RAT	P46892 rattus norv
29	54	15.6	494	1 VABR_PLAFA	O25691 plasmodium
30	54	15.6	682	1 KPCL_HUMAN	P24723 homo sapien
31	54	15.6	983	1 YOG4_CABEL	O09499 caenorhabdi
32	53.5	15.5	310	1 Y160_BUCAP	O8K934 buchiera ap
33	53.5	15.5	394	1 Y129_CHLUP	O9JRX2 chlamydia p

34	53.5	15.5	609	1 SYD_CAUCR	O94734 caulobacter
35	53.5	15.5	705	1 Y450_RHISN	P56566 rhizobium s
36	53	15.3	436	1 KP58_HUMAN	P21127 homo sapien
37	53	15.3	549	1 SYR_ARCFU	O29368 archaeoglob
38	53	15.3	695	1 TKT2_RHIME	P56900 rhizobium m
39	52.5	15.2	74	1 SOM1_YEAST	O05676 saccharomyc
40	52.5	15.2	282	1 HKCC_HUMAN	P31275 homo sapien
41	52.5	15.2	317	1 GGH_RAT	O62867 rattus norv
42	52.5	15.2	350	1 FLNA_TREPA	P18193 treponema p
43	52.5	15.2	360	1 DCAM_SPIOL	P46255 spinacia ol
44	52.5	15.2	366	1 YBAB_BACCI	P48843 bacillus ci
45	52.5	15.2	372	1 OPBD_MOUSE	P32300 mus musculu

ALIGNMENTS

RESULT 1
ID CAIH_HUMAN STANDARD; PRT; 1516 AA.
AC P39060; O9UK38; O9Y6Q7; O9Y6Q8;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Collagen alpha 1(XVII) chain precursor (Contains: Endostatin).
GN COL18A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98164096; PubMed=9503365;
RA Saarela J., Ylikarppa R., Rehn M., Purmonen S., Pihlajaniemi T.,
RT "Complete primary structure of two variant forms of human type XVIII
RT collagen and tissue-specific differences in the expression of the
RT corresponding transcripts."
RL Matrix Biol. 16:319-328(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hactori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
RA Soeda E., Onki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shinomiya A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloedker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Mehrezyer S., Borzym K., Gardiner K., Niretic D., Francis F.,
RA Lehrs H., Reinhardt R., Yaspo M.-L.,
RT "The DNA sequence of human chromosome 21."
RL Nature 405:311-319(2000).
RN [3]
RP SEQUENCE OF 834-1516 FROM N.A.
RX MEDLINE=94245237; PubMed=8186291;
RA Oh S.P., Warman M.L., Seidlin M.F., Cheng S., Knoll J.H., Timmons S.,
RA Olsen B.R.,
RT "Cloning of cDNA and genomic DNA encoding human type XVIII collagen
RT and localization of the alpha 1(XVII) collagen gene to mouse
RT chromosome 10 and human chromosome 21."
RL Genomics 19:494-499(1994).
RN [4]
RP SEQUENCE OF 1334-1516 FROM N.A.
RX TISSUE=Placenta;
RA Zhi-Yong H., Biao L., Wei-Jie Z., Xiang-Fu W.,
RT "Cloning and expression of human endostatin gene in Escherichia
RT coli."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP INVOLVEMENT IN KNOBLOCH SYNDROME.

DR	SMART; SMO0282; LamG; 1.
DR	SMART; SMO0210; TSPN; 1.
KW	Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW	Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing;
KW	Polyomorphism; 3d-structure.
FT	SIGNAL 1 23
FT	CHAIN 24 1516
FT	CHAIN 1334 1516
FT	DOMAIN 24 516
FT	DOMAIN 517 550
FT	DOMAIN 551 560
FT	DOMAIN 561 640
FT	DOMAIN 641 664
FT	DOMAIN 665 786
FT	DOMAIN 787 809
FT	DOMAIN 810 892
FT	DOMAIN 893 906
FT	DOMAIN 907 948
FT	DOMAIN 949 961
FT	DOMAIN 962 1034
FT	DOMAIN 1035 1044
FT	DOMAIN 1045 1077
FT	DOMAIN 1078 1089
FT	DOMAIN 1090 1111
FT	DOMAIN 1112 1118
FT	DOMAIN 1119 1173
FT	DOMAIN 1174 1186
FT	DOMAIN 1187 1204
FT	DOMAIN 1205 1516
FT	CARBOHYD 68
FT	CARBOHYD 129 129
FT	CARBOHYD 164 164
FT	CARBOHYD 691 691
FT	CARBOHYD 1329 1329
FT	DISULFID 1366 1506
FT	DISULFID 1468 1498
FT	SITE 1095 1097
FT	VARSPLIC 1 180
FT	VARSPLIC 181 215
FT	VARIANT 1437 1437
FT	CONFLICT 428 428
FT	CONFLICT 841 841
FT	CONFLICT 877 877
FT	CONFLICT 886 886
FT	CONFLICT 912 912
FT	CONFLICT 933 933
FT	CONFLICT 975 975
FT	CONFLICT 1064 1064
FT	CONFLICT 1084 1084
FT	CONFLICT 1120 1120
FT	CONFLICT 1123 1123
FT	CONFLICT 1126 1126
FT	CONFLICT 1206 1206
FT	CONFLICT 1304 1304
FT	CONFLICT 1314 1314
FT	CONFLICT 1323 1324
FT	CONFLICT 1443 1443
FT	CONFLICT 1483 1483
SO	SEQUENCE 1516 AA; 153840 MW; 3C70F29A4476EB76 CRC64;
Query Match	Similarity 100.0%; Score 346; DB 1; Length 1516;
Best Local	Similarity 100.0%; Pctd. No. 8.1e-33;
Matches 63;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1397 ADRAAVPIVULKDELLFPPSWBALFSGSEGPLKPGARIISPDGKDVLRHPTMPQKSVWHS 1456

OY 61 DPN 63

DB 1457 DPN 1459

RESULT 2

CAH_MOUSE STANDARD; PRT; 1527 AA.

AC P39061; 061437; 062002;

DT 01-FEB-1995 (Rel. 31, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 15-SEP-2003 (Rel. 42, Last annotation update)

GN COL18A1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1] SEQUENCE FROM N.A. (ISOFORM SHORT).

RP STRAIN=BALB/C; TISSUE=Liver;

RX MEDLINE=94245707; PubMed=818673;

RA Rehn M.V., Hintikka E., Pihlajaniemi T.;

RT "Primary structure of the alpha 1 chain of mouse type XVIII collagen, partial structure of the corresponding gene, and comparison of the alpha 1(XVIII) chain with its homologue, the alpha 1(XV) collagen chain."

RT J. Biol. Chem. 269:13929-13935(1994).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).

RA Rehn M., Hintikka E., Pihlajaniemi T.;

RT "Characterization of the mouse gene for the alpha-1 chain of type XVIII collagen (COL18A1) reveals that the three variant N-terminal polypeptide forms are transcribed from two widely separated promoters."

RT Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 213-1140 FROM N.A. (ISOFORM SHORT).

RX MEDLINE=94240112; PubMed=8183894;

RA Rehn M.V., Pihlajaniemi T.;

RT "Alpha 1(XVIII), a collagen chain with frequent interruptions in the collagenous sequence, a distinct tissue distribution, and homology with type XV collagen."

RT Proc. Natl. Acad. Sci. U.S.A. 91:4234-4238(1994).

RN [4]

RP SEQUENCE OF 240-1527 FROM N.A.

RX TISSUE=Liver;

RX MEDLINE=94240111; PubMed=8183893;

RA Oh S.P., Kamagata Y., Muragaki Y., Timmons S., Ooshima A., Olsen B.R.;

RT "Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xaa-Yaa repeats identify a distinct family of collagenous proteins."

RT Proc. Natl. Acad. Sci. U.S.A. 91:4229-4233(1994).

RN [5]

RP CHARACTERIZATION OF ENDOSTATIN AND PARTIAL SEQUENCE.

RX MEDLINE=97160848; PubMed=9008168;

RA O'Reilly M.S., Boehm T., Shing Y., Fukai N., Vasios G., Lane W.S., Flynn E., Birkhead J.R., Olsen B.R., Folkman J.;

RT "Endostatin: an endogenous inhibitor of angiogenesis and tumor growth."

RT Cell 86:277-285(1997).

RN [6]

RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF ENDOSTATIN.

RX MEDLINE=98169382; PubMed=9501087;

RA Hohenester E., Sasaki T., Olsen B.R., Timpl R.;

RT "Crystal structure of the angiogenesis inhibitor endostatin at 1.5-A resolution."

RT EMBO J. 17:1656-1664(1998).

CC - FUNCTION: ENDOSTATIN POTENTIALLY INHIBITS ENDOTHELIAL CELL PROLIFERATION AND ANGIOGENESIS. MAY INHIBIT ANGIOGENESIS BY BINDING TO THE HEPARAN SULPHATE PROTEOGLYCANS INVOLVED IN GROWTH

CC FACTOR SIGNALING.

CC - ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=Long;

CC IsoId=P39061-1; Sequence=Displayed;

CC Name=Short;

CC IsoId=P39061-2; Sequence=VSP 001157, VSP 001158;

CC - PTM: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated in some or all of the chains.

CC - SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH INTERRUPTED HELICES (FACIT) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; I16898; AAA37434.1; -

DR EMBL; U03714; AAC20657.1; -

DR EMBL; U03715; AAC52901.1; JOINED.

DR EMBL; U34606; AAC52901.1; JOINED.

DR EMBL; U34608; AAC52901.1; JOINED.

DR EMBL; U34609; AAC52901.1; JOINED.

DR EMBL; U34610; AAC52901.1; JOINED.

DR EMBL; U34611; AAC52901.1; JOINED.

DR EMBL; U34612; AAC52901.1; JOINED.

DR EMBL; U34613; AAC52901.1; JOINED.

DR EMBL; U03716; AAC52901.1; JOINED.

DR EMBL; U03718; AAC52901.1; JOINED.

DR EMBL; U03715; AAC52902.1; JOINED.

DR EMBL; U34608; AAC52902.1; JOINED.

DR EMBL; U34609; AAC52902.1; JOINED.

DR EMBL; U34610; AAC52902.1; JOINED.

DR EMBL; U34611; AAC52902.1; JOINED.

DR EMBL; U03716; AAC52902.1; JOINED.

DR EMBL; U03718; AAC52902.1; JOINED.

DR EMBL; U11636; AAC52178.1; -

DR EMBL; L22545; AAA19787.1; -

DR PIR; A56101; A56101.

DR PDB; 1KOE; 16-FEB-99.

DR PDB; 1DY0; 11-APR-00.

DR PDB; 1DY1; 21-JAN-01.

DR MGD; MGI:88451; COL18A1.

DR GO; GO:0001525; P:angiogenesis; IMP.

DR InterPro; IPR000087; Collagen.

DR InterPro; IPR001791; Laminin_G.

DR InterPro; IPR003129; TSPN.

DR Pfam; PF01391; Collagen; 8.

DR Pfam; PF02210; TSPN; 1.

DR ProDom; PD000007; C1g_helix; 1.

DR SMART; SMO0282; Lamg; 1.

DR SMART; SMO0210; TSPN; 1.

DR Extracellular matrix; Connective tissue; Repeat; Hydroxylation; Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing; 3D-structure.

KM 3D-structure.

FT SIGNAL 1 26

FT CHAIN 27 1527

FT CHAIN 1344 1527

FT DOMAIN 27 538

FT DOMAIN 539 565

FT DOMAIN 566 575

FT DOMAIN 576 649

FT DOMAIN 650 673

FT DOMAIN 674 795

FT DOMAIN 796 818

FT DOMAIN 819 901

FT DOMAIN 902 915

POTENTIAL.

COLLAGEN ALPHA 1(XVIII) CHAIN.

ENDOSTATIN.

NONHELICAL REGION 1 (NC1).

TRIPLE-HELICAL REGION 1 (COL1).

NONHELICAL REGION 2 (NC2).

TRIPLE-HELICAL REGION 2 (COL2).

NONHELICAL REGION 3 (NC3).

TRIPLE-HELICAL REGION 3 (COL3).

NONHELICAL REGION 4 (NC4).

TRIPLE-HELICAL REGION 4 (COL4).

NONHELICAL REGION 5 (NC5).

```

FT DOMAIN 916 957 TRIPLE-HELICAL REGION 5 (COL5).
FT DOMAIN 958 970 NONHELIICAL REGION 6 (NC6).
FT DOMAIN 971 1043 TRIPLE-HELICAL REGION 6 (COL6).
FT DOMAIN 1044 1053 NONHELIICAL REGION 7 (NC7).
FT DOMAIN 1054 1086 TRIPLE-HELICAL REGION 7 (COL7).
FT DOMAIN 1087 1098 NONHELIICAL REGION 8 (NC8).
FT DOMAIN 1099 1122 TRIPLE-HELICAL REGION 8 (COL8).
FT DOMAIN 1123 1129 NONHELIICAL REGION 9 (NC9).
FT DOMAIN 1130 1181 TRIPLE-HELICAL REGION 9 (COL9).
FT DOMAIN 1182 1194 NONHELIICAL REGION 10 (NC10).
FT DOMAIN 1195 1212 TRIPLE-HELICAL REGION 10 (COL10).
FT DOMAIN 1213 1527 NONHELIICAL REGION 11 (NC11).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 700 700 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 1376 1516
FT DISULFID 1478 1508
FT SITE 1104 1106
FT VARSPPLIC 1 212
FT VARSPPLIC 213 238

CELL ATTACHMENT SITE (POTENTIAL).
Missing (in isoform short).
/FTid=VSP 001157.
AAPTOLPPQSNIOAPLGRSPAPDF -> MAPRMLDLVL
TSVLILVARVMAE (in isoform short).
/FTid=VSP 001158.
P -> L (IN REF. 4).
P -> F (IN REF. 4).
A -> R (IN REF. 4).
R -> P (IN REF. 4).
P -> L (IN REF. 4).
L -> F (IN REF. 4).
L -> V (IN REF. 4).

FT CONFLICT 900 900
FT CONFLICT 947 947
FT CONFLICT 964 964
FT CONFLICT 1157 1157
FT CONFLICT 1266 1266
FT CONFLICT 1276 1276
FT CONFLICT 1437 1437
FT STRAND 1353 1357
FT STRAND 1362 1362
FT HELIX 1368 1382
FT TURN 1383 1384
FT STRAND 1389 1391
FT STRAND 1395 1395
FT TURN 1396 1397
FT STRAND 1398 1398
FT HELIX 1400 1403
FT HELIX 1406 1408
FT TURN 1409 1411
FT STRAND 1414 1415
FT TURN 1417 1418
FT STRAND 1421 1423
FT HELIX 1426 1429
FT TURN 1430 1430
FT STRAND 1436 1436
FT TURN 1439 1440
FT STRAND 1444 1445
FT TURN 1446 1447
FT STRAND 1450 1450
FT TURN 1451 1453
FT TURN 1455 1456
FT STRAND 1461 1463
FT STRAND 1466 1466
FT TURN 1468 1469
FT STRAND 1472 1472
FT TURN 1474 1475
FT STRAND 1477 1477
FT HELIX 1478 1481
FT TURN 1482 1482
FT TURN 1486 1487
FT STRAND 1489 1494
FT HELIX 1495 1497

```

Query Match 85.0%; Score 294; DB 1; Length 1527;
 Best Local Similarity 81.0%; Pred. No. 1.1e-26;
 Matches 51; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

```

Qy 1 ADRAAVPIVVKKEELFSPSEALFSGSEGLKPGARIFSDGQOVLRHPTWPKQSVWHS 60
Db 1407 ADRAAVPIVVKKEELFSPSEALFSGSEGLKPGARIFSDGQOVLRHPTWPKQSVWHS 1466
Qy 61 DPN 63

```

```

Db 1467 DPS 1469
||:
RESULT 3
ID CALE HUMAN STANDARD; PRT, 1388 AA.
AC P39059;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Collagen alpha 1(XV) chain precursor.
GN COL15A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=umbilical cord;
RX MEDLINE=941148920; PubMed=8106446;
RA Kivirikko S., Heinämäki P., Rehn M.V., Honkanen N., Myers J.C.,
RA Pihlajaniemi T.;
RT "Primary structure of the alpha 1 chain of human type XV collagen and
RT exon-intron organization in the 3' region of the corresponding
RT gene."
RL J. Biol. Chem. 269:4773-4779(1994).
RN [2]
RP SEQUENCE OF 1-569 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=941140817; PubMed=8307960;
RA Murgaki Y., Abe N., Ninomiya Y., Olsen B.R., Ooshima A.;
RT "The human alpha 1(XV) collagen chain contains a large amino-terminal
RT non-triple helical domain with a tandem repeat structure and homology
RT to alpha 1(XVII) collagen."
RL J. Biol. Chem. 269:4042-4046(1994).
RN [3]
RP SEQUENCE OF 544-1252 FROM N.A.
RX MEDLINE=93066196; PubMed=1279671;
RA Myers J.C., Kivirikko S., Gordon M.K., Dion A.S., Pihlajaniemi T.;
RT "Identification of a previously unknown human collagen chain, alpha
RT 1(XV), characterized by extensive interruptions in the triple-helical
RT region."
RL Proc. Natl. Acad. Sci. U.S.A. 89:10144-10148(1992).
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN INTERNAL ORGANS
CC SUCH AS ADRENAL GLAND, PANCREAS AND KIDNEY.
CC -1- PTM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -1- SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH
CC INTERRUPTED HELICES (FACIT) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL, L25286; AAA58429.1; .
CC EMBL, D21230; BAA04762.1; .
CC EMBL, L01697; -; NOT_ANNOTATED_CDS.
CC PIR, A53317; A53317.
CC HSPF, P39061; IKOE.
CC Genew; HGNC:2192; COL15A1.
CC MIM; 120325; .
CC InterPro; IPR000087; Collagen.
CC InterPro; IPR001791; Laminin_G.
CC InterPro; IPR003129; TSPN.
CC Pfam; PF01391; Collagen_5.
CC Pfam; PF02210; TSPN_1.
CC SMART; SM00282; LamG_1.
CC SMART; SM00210; TSPN_1.

```


FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: A061265; AAC16713.1; -
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP BINDING; 1.
 KW Ligase; Multigene family.
 SQ SEQUENCE 683 AA; 76029 MW; 7AC88764A97966DB CRC64;

Query Match 17.6%; Score 61; DB 1; Length 683;
 Best Local Similarity 27.1%; Pred. No. 16;
 Matches 16; Conservative 8; Mismatches 13; Indels 22; Gaps 3;

QY 25 SGSEGLKPKGARIFFSGDKVLRHPT-----WPG-KSVHGGSD 61
 DB 460 AGGATPMKFGAAMPFFFGIDLVDPDTGIEGTGEHAGVLAIKRPMSPARTIMKND 518

RESULT 6
 ID PRTS_SERMA STANDARD; PRT; 1045 AA.
 AC P09489;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Extracellular serine protease precursor (EC 3.4.21.-).
 OS *Serratia marcescens*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Serratia.
 OX NCBI_TaxId=615;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-35 AND 407-408.
 RC STRAIN=IFO 3046;
 RA MEDLINE=66223815; PubMed3011754;
 RA Yanagida N., Uozumi T., Beppu T.;
 RT "Specific excretion of Serratia marcescens protease through the outer
 RT membrane of Escherichia coli.";
 RL J. Bacteriol. 166:937-944 (1986).
 RP PARTIAL SEQUENCE, AND PROCESSING.
 RX MEDLINE=92348352; PubMed1639760;
 RA Shikata S., Shimada K., Kataoka H., Hornouchi S., Beppu T.;
 RT "Detection of large COOH-terminal domains processed from the
 RT precursor of Serratia marcescens serine protease in the outer
 RT membrane of Escherichia coli.";
 RL J. Biochem. 111:627-632 (1992).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M13469; AA26572.1; -
 DR PIR: A29840; A29840.
 DR HSR: Q99405; IMPT.
 DR MEROPS: S08.094; -
 DR InterPro: IPR006315; Autotransport.
 DR InterPro: IPR005546; Autotransporter.
 DR InterPro: IPR000209; peptidase_S8.
 DR Pfam: PF03797; Autotransporter; 1.

DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR00723; SUBTILISTIN.
 DR TIGRFAMs: TIGR01414; autotrans_bar1; 1.
 DR PROSITE: PS00136; SUBTILASE ASP; FALSE_NEG.
 DR PROSITE: PS00137; SUBTILASE HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Hydrolase; Serine protease; zymogen; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 645 EXTRACELLULAR SERINE PROTEASE.
 FT PROPEP 646 1045
 FT ACT SITE 76 76 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 112 112 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 341 341 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 1045 AA; 112345 MW; 4924E450E4FF179C CRC64;

Query Match 17.6%; Score 61; DB 1; Length 1045;
 Best Local Similarity 34.9%; Pred. No. 25;
 Matches 15; Conservative 4; Mismatches 24; Indels 0; Gaps 0;

QY 14 ELLFSSWEALFSGSGPLKPKGARIFFSGDKVLRHPTVPQSV 56
 DB 449 ECSFDSMSNDISGHGGLTKTGAGTLALGNNTYRDTWVKQV 491

RESULT 7
 ID TKT1_RHIME STANDARD; PRT; 694 AA.
 AC P58333;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transketolase (EC 2.2.1.1) (TK).
 GN CBRT OR RB0193 OR SMD2020.
 OS *Rhizobium meliloti* (Sinorhizobium meliloti).
 OC Plasmid pSymB (megaplasmid 2).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxId=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396508; PubMed11481431;
 RA Finan T.M., Weidner S., Wong K., Buhrmaster J., Chain P.,
 RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
 RA Golding B., Pecher A.;
 RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the
 RT N2-fixing endosymbiont Sinorhizobium meliloti.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894 (2001).
 CC -1- CATALYTIC ACTIVITY: Sedoheptulose 7-phosphate + D-glyceraldehyde
 CC 3-phosphate = D-ribose 5-phosphate + D-xylulose 5-phosphate.
 CC -1- COFACTOR: THIAMINE PYROPHOSPHATE (By similarity).
 CC -1- PATHWAY: Calvin cycle.
 CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AL603642; CAC68593.1; -
 DR PIR: A95866; A95866.
 DR InterPro: IPR005478; BactTransketolase.
 DR InterPro: IPR005476; Transketolase_C.
 DR InterPro: IPR005475; Transketolase_N.
 DR InterPro: IPR005474; Transketolase_CR.
 DR Pfam: PF03779; transket_pyr; 1.
 DR Pfam: PF00456; transketolase; 1.
 DR Pfam: PF02780; transketolase_C; 1.
 DR TIGRFAMs: TIGR00232; tkLase_bact; 1.
 DR PROSITE: PS00801; TRANSKETOLASE_1; 1.


```

CC (POTENTIAL).
CC - DOMAIN: Lacks the zinc protease domain of other ftsH proteins.
CC Also much longer in both the N- and C-termini.
CC - SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB001684; BAA57905.1; -.
CC EMBL; AB001684; BAA57906.1; ALT_INIT.
CC PIR; T07258; T07258.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003959; AAA_ATPase_cent.
CC InterPro; IPR003960; AAA_gub.
CC InterPro; IPR006642; Peptidase_M41.
CC Pfam; PF00004; AAA; 1.
CC Pfam; PF01434; Peptidase_M41; 1.
CC SMART; SM00382; AAA; 1.
CC PROSITE; PS00674; AAA; 1.
CC Cell division; ATP-binding; Transmembrane; Chloroplast.
CC KW TRANSMEM 48 68 POTENTIAL.
CC FT TRANSMEM 896 916 POTENTIAL.
CC FT TRANSMEM 973 993 POTENTIAL.
CC FT TRANSMEM 1021 1041 POTENTIAL.
CC SQ SEQUENCE 1720 AA; 197172 MW; 0F1EA926B799D5B8 CRC64;

Query Match 16.9%; Score 58.5; DB 1; Length 1720;
Best Local Similarity 26.7%; Pred. No. 86;
Matches 16; Conservative 14; Mismatches 21; Indels 9; Gaps 4;

QY 9 VNLKDELLF--PSEWALFSGSGEPLKP-----GARIFSGDKDVLNHPW--PKSVWHS 60
DB 1480 IGFEDDIRYSSPMWQDVS-AEMEFKPKPKGSRILYDERTSNPEWVPDERVYHS 1538

RESULT 10
ACSA_SCHPO STANDARD; PRT; 662 AA.
ID ACSA_SCHPO
AC P78773;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable acetyl-coenzyme A synthetase (EC 6.2.1.1) (Acetate--CoA
DE 1ligase) (Acyl-activating enzyme).
GN SPC191.02C OR SPC417.14C.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4836;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream N.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jasele K.,
RA James K., Jones L., Jones M., Leach S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Nibbel D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens S., Taylor K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grynoperez B.,
RA Weljens I., Vanetere E., Rieger M., Schaefer M., Mueller-Auer S.,

```

```

RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
RA Goffeau A., Cadiou E., Dreano S., Gloux S., Lelure V., Mottier S.,
RA Gilbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucan M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen O., Potashkin J.,
RA Shpakovski G.V., Usero D., Barrett B.G., Nurse P.,
RA "The genome sequence of Schizosaccharomyces pombe."
RA Nature 415:871-880(2002).
RN [12]
RP SEQUENCE OF 97-662 FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshioke S., Kato K., Nakai K., Okayama H., Nojima H.,
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT cDNA."
RT DNA Res. 4:363-369(1997).
CC -1- CATALYTIC ACTIVITY: ATP + acetate + CoA = AMP + diphosphate +
CC acetyl-CoA.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL049644; CAB41048.1; -.
CC EMBL; AL035076; CAA22660.1; -.
CC EMBL; D89121; BAA13783.1; -.
CC PIR; T41215; T41215.
CC GenDB; Spombe; SPC191.02C; -.
CC InterPro; IPR000873; AMP-bind.
CC Pfam; PF00501; AMP-binding; 1.
CC PROSITE; PS0455; AMP_BINDING; 1.
CC KW 1ligase.
CC SQ SEQUENCE 662 AA; 73035 MW; 02FD119ABB622583 CRC64;

Query Match 16.8%; Score 58; DB 1; Length 662;
Best Local Similarity 28.8%; Pred. No. 35;
Matches 17; Conservative 6; Mismatches 12; Indels 24; Gaps 4;

QY 26 GSEGLPKGARGIFSF-----DGKDV-----LRHPTWPO--KSVWHS 61
DB 435 GVTMTKSGSATLPFGIDAVIIDLTKITIEGNDVEGLAIRSF-WPSARTVRGHD 492

RESULT 11
FXC1_HUMAN STANDARD; PRT; 553 AA.
ID FXC1_HUMAN
AC O12948; Q9BYM1; Q9NUE5; Q9UDD0; Q9UP06;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Forkhead box protein C1 (Forkhead-related protein FKHL7) (Forkhead-
DE related transcription factor 3) (FKHRC-3).
GN FOXO1 OR FKHL7 OR FKRC3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS SER-112; MET-126 AND LEU-131.
RC MEDLINE=96282091; PubMed=9620769;
RX Nishimura D.Y., Swideraki R.E., Alward W.L.M., Seaby C.C.,
RA Pacil S.R., Bennett S.R., Kanis A.B., Gaetflier J.M., Stone E.M.,
RA Sheffield V.C.;

```

[illegible]

binding domain.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PKC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D90242; AAA14288.1; -
 DR PIR: A23690; A23690.
 DR HSSP: P28867; IPTQ.
 DR MGD: MG197600; PKCh.
 DR InterPro: IPR000008; C2.
 DR InterPro: IPR002219; DAG_PE-bind.
 DR InterPro: IPR000961; PKinase_C.
 DR InterPro: IPR000719; Prot_Kinase.
 DR InterPro: IPR002290; Ser_Thr_PKinase.
 DR Pfam: PF00168; C2; 1.
 DR Pfam: PF00130; DAG_PE-bind; 2.
 DR Pfam: PF00069; PKinase; 1.
 DR Pfam: PF00433; PKinase; C; 1.
 DR PRINTS: PR00006; DAGPEDOMAIN.
 DR ProDom: PD000001; Prot_Kinase; 1.
 DR SMART: SM00109; C1; 2.
 DR SMART: SM00230; C2; 1.
 DR SMART: SM00133; S_TK_X; 1.
 DR SMART: SM00320; S_TK_X; 1.
 DR PROSITE: PSS0004; C2 DOMAIN 2; 1.
 DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; 2.
 DR PROSITE: PS00479; DAG_PE_BIND_DOM_2; 2.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
 DR APP-binding: Transferase; Serine/threonine-protein kinase;
 KW Phorbol-ester binding; Zinc; Repeat.
 KW DOMAIN 12 102
 FT DOMAIN 172 222 PHORBOL-ESTER AND DAG BINDING 1.
 FT DOMAIN 246 295 PHORBOL-ESTER AND DAG BINDING 2.
 FT DOMAIN 355 614 PROTEIN KINASE.
 FT NP_BIND 361 369 ATP (BY SIMILARITY).
 FT BINDING 384 384 ATP (BY SIMILARITY).
 FT ACT_SITE 479 479 BY SIMILARITY.
 SQ SEQUENCE 683 AA; 77973 MW; C7DB338A9F95F576 CRC64;
 Query Match 16.5%; Score 57; DB 1; Length 683;
 Best Local Similarity 31.6%; Pred. No. 47;
 Matches 18; Conservative 8; Mismatches 15; Indels 16; Gaps 4;
 QY 11 LKDELFPSEWALFSGSEGLK-----PGARIFSF--DGKVLRRP-----TWPG 53
 DB 567 LNDVVVPTV--LHEDARGLIKSFWMTNPTMRGLSLQGGEHEILRRPFKEIDWAG 621
 RESULT 13
 TRPG ASPNG STANDARD; PRT; 770 AA.
 AC P05328;
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Anthranilate synthase component II (EC 4.1.3.27) [includes: Glutamine
 DE amidotransferase; Indole-3-glycerol phosphate synthase (EC 4.1.1.48)
 DE (IGPS); N-(5'-phosphoribosyl)anthranilate isomerase (EC 5.3.1.24)
 DE (PRAI)].
 GN TRC.
 OS *Aspergillus niger*.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; *Aspergillus*.

NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=401;
 RA MEDLINE=88223483; PubMed=2836085;
 RA Kos T., Kuijvenhoven A., Hoesing H.G.M., Pouwels P.H.,
 RA van den Hondel C.A.M.J.J.;
 RT "Nucleotide sequence of the *Aspergillus niger* trpC gene: structural
 RT relationship with analogous genes of other organisms";
 RN Curr. Genet. 13:137-144(1988).
 RN [2]
 RP SEQUENCE OF 1-69 AND 392-433 FROM N.A.
 RX MEDLINE=86137391; PubMed=2936550;
 RA Kos A., Kuijvenhoven J., Wernars K., Bos C.J., van den Broek H.W.J.,
 RA Pouwels P.H., van den Hondel C.A.M.J.J.;
 RT "Isolation and characterization of the *Aspergillus niger* trpC gene";
 RL Gene 39:231-238(1985).
 CC -1- FUNCTION: TRIFUNCTIONAL ENZYME BEARING THE GLN AMIDOTRANSFERASE
 CC (GATASE) DOMAIN OF ANTHRANILATE SYNTHASE, INDOLE-GLYCEROLPHOSPHATE
 CC SYNTHASE, AND PHOSPHORIBOSYLANTHRANILATE ISOMERASE ACTIVITIES.
 CC -1- CATALYTIC ACTIVITY: N-(5'-phospho-beta-D-ribose)-anthranilate = 1-
 CC (2-carboxyphenylamino)-1-deoxy-D-ribulose 5-phosphate
 CC -1- CATALYTIC ACTIVITY: 1-(2-carboxyphenylamino)-1-deoxy-D-ribulose 5-
 CC phosphate = 1-(indol-3-yl)glycerol 3-phosphate + CO(2) + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Chorismate + L-glutamine = anthranilate +
 CC pyruvate + L-glutamate.
 CC -1- PATHWAY: Tryptophan biosynthesis; first step.
 CC -1- PATHWAY: Tryptophan biosynthesis; third step.
 CC -1- PATHWAY: Tryptophan biosynthesis; fourth step.
 CC -1- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X07071; CAA30107.1; -
 DR EMBL: M14403; AAA32710.1; -
 DR EMBL: M14403; AAA32709.1; -
 DR PIR: S00643; S00643.
 DR HSSP: Q06129; IQDL.
 DR InterPro: IPR006220; Anth synthII.
 DR InterPro: IPR001317; CPS_Gatase.
 DR InterPro: IPR003009; FMN_enzyme.
 DR InterPro: IPR000991; Gatase_1.
 DR InterPro: IPR001468; IGPS_1.
 DR InterPro: IPR006221; TrpG_papa.
 DR Pfam: PF00117; Gatase; 1.
 DR Pfam: PF00218; IGPS; 1.
 DR Pfam: PF00697; PRAI; 1.
 DR PRINTS: PR00097; ANTSYNTHASEII.
 DR PRINTS: PR00099; CPSGATASE.
 DR PRINTS: PR00096; GATASE.
 DR ProDom: PD001511; IGPS; 1.
 DR TIGRFAMs: TIGR00366; trpG_papa; 1.
 DR PROSITE: PS00442; GATASE_TYPE_1; 1.
 DR PROSITE: PS00614; IGPS; 1.
 KW Tryptophan biosynthesis; Isomerase; Lyase; Multifunctional enzyme;
 KW decarboxylase; Transferase; Glutamine amidotransferase;
 FT DOMAIN 25 219
 FT DOMAIN 255 519
 FT DOMAIN 535 770
 FT ACT_SITE 104 104
 FT ACT_SITE 199 199
 FT ACT_SITE 201 201
 SQ SEQUENCE 770 AA; 82909 MW; 36D8DE5B23097012 CRC64;
 Query Match 16.3%; Score 56.5; DB 1; Length 770;

Best Local Similarity 30.6%; Pred. No. 62;
Matches 15; Conservative 7; Mismatches 26; Indels 1; Gaps 1;

Qy 2 DRAAPVNLKDELFPSEWALFSGSGPLKPGARIFSPDGKVLPHPT 50
Db 572 DDVALPISQVVRSTPKPASQALHTSOEPKATSVYFDHSAK-ILRHPS 619

RESULT 14

IP67_MYCTU
ID TYPE7_MYCTU STANDARD; PRT; 884 AA.
AC OS0654; OS0731;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein RV2567.
GN RV2567 OR MT2643 OR MTCY227.34C OR MTCY9C4.01C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxId=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Bismeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornby S., Jacobs K., Kiroch A., McLean A., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: SOME, TO M.TUBERCULOSIS RV2411C AND SYNECHOCYSTIS PCC
CC 6803 SLI0335.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL: Z77250; CAB01053.1;
CC EMBL: AE007098; AAK46956.1;
CC PIR: C70729; C70729.
CC TIGR: MT2643;
DR Tuberculin; RV2567;
DR Pfam: PF04168; DUF403; 1.
DR Pfam: PF04169; DUF404; 1.
DR Pfam: PF04174; DUF407; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 645 645 O -> R (IN REF. 2).
SQ SEQUENCE 884 AA; 95448 MW; 95023A4D2EDB365 CRC64;

Query Match 16.3%; Score 56.5; DB 1; Length 884;
Best Local Similarity 30.2%; Pred. No. 72;
Matches 16; Conservative 11; Mismatches 13; Indels 13; Gaps 2;

Qy 23 LFGSGRGLKPGARIFSPDG-----KDVLRHPT-----PQKSWHGSDD 62
Db 620 LMSLTVPDPSPGLVQSVGLALAAQAVRDQSLNDTMYLVANVERAVERKSDP 672

RESULT 15

SYM_METUA
ID SYM_METUA STANDARD; PRT; 651 AA.
AC OS8659;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Methionyl-L-cysteine synthetase (EC 6.1.1.10) (Methionine--cysteine ligase)
DE (MetRS)
GN METG OR MJ1263.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxId=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Funtman J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Huzar M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).

CC -1- FUNCTION: Is required not only for elongation of protein synthesis
CC but also for the initiation of all mRNA translation through
CC initiator tRNA (fMet) aminoacylation (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
CC diphosphate + L-methionyl-tRNA(Met).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC MetG subfamily 1.
CC -1- SIMILARITY: Contains 1 tRNA-binding domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL: U67567; AAB99269.1;
CC PIR: F64457; F64457.
CC HSSP: P00959; IMEA.
CC TIGR: MJ1263;
DR HAMAP: MF_00098; fused. 1.
DR InterPro: IPR004495; MetG_Cterm.
DR InterPro: IPR002300; tRNA-synt_1a.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002304; tRNA-synt_met.
DR InterPro: IPR002547; tRNA-bind.
DR Pfam: PF00133; tRNA-synt_1; 1.
DR Pfam: PF01588; tRNA-bind; 1.
DR PIRSF: PIRSF001528; MetRS_dimerising; 1.
DR PRINTS: PR01041; TRNASTYTHMET.
DR TIGRFAMs: TIGR00398; metG; 4.
DR TIGRFAMs: TIGR00399; metG_Cterm; 1.
DR PROSITE: PS00178; AA tRNA_LIGASE_I; FALSE_NEG.
DR PROSITE: PS50866; TRBD; 1.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: September 10, 2003, 14:09:48 ; Search time 21 Seconds

(without alignments)
288.506 Million cell updates/sec

Title: US-09-822-540A-1

Perfect score: 346

Sequence: 1 ADRAAVPIVNLKDELLFSPW.....DVLRHPTWPKSVHSGDPN 63

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	346	100.0	684	2 AS5019	collagen alpha 1(X
2	294	85.0	1774	2 B56101	collagen alpha 1(X
3	191	84.1	1315	2 AS6101	collagen alpha 1(X
4	193	55.8	1388	2 AS3317	collagen alpha 1(X
5	140	40.5	650	2 T22002	hypothetical prote
6	67.5	19.5	614	2 S27962	modulator recognit
7	64	18.5	644	2 A72519	probable 2-oxoacid
8	63.5	18.4	266	2 S22511	chlorophyll a/b-bi
9	63	18.2	336	2 T17408	RNA (adenine-N6-)
10	62.5	18.1	467	2 T26705	hypothetical prote
11	62.5	18.1	620	2 F68797	conserved hypochet
12	62	17.9	419	2 T25565	hypothetical prote
13	61	17.6	1045	2 A29840	serine proteinase
14	60	17.3	694	2 A95866	probable transkeo
15	60	17.3	984	2 T44496	cellulose 1,4-beta
16	59.5	17.2	268	2 H75614	hypothetical prote
17	59.5	17.2	362	2 S24434	class I histocompa
18	59	17.1	303	2 T41056	conserved hypochet
19	58.5	16.9	305	2 S07115	class I histocompa
20	58.5	16.9	354	2 I80170	class I histocompa
21	58.5	16.9	481	2 T21311	hypothetical prote
22	58.5	16.9	1184	2 D63387	probable protein p
23	58.5	16.9	1720	2 T07258	cell division prot
24	58	16.8	461	2 A54024	protein kinase (EC
25	58	16.8	662	2 T41215	probable acetate-C
26	58	16.8	777	2 AD0982	biclin sulfoxide r
27	58	16.8	2055	2 T30259	multiple PDZ domai
28	57.5	16.6	506	2 H64618	sigma-54 interacti
29	57.5	16.6	732	2 AF0439	probable Rhs acces

30	57.5	16.6	800	2 AG0438	probable Rhs acces
31	57	16.5	335	2 A23174	hypothetical prote
32	57	16.5	683	2 A23690	protein kinase (EC
33	57	16.5	1116	2 B88612	protein Y6A2A.2 (
34	57	16.5	1238	1 JC5573	copper-transporin
35	57	16.5	2054	2 T46612	multi PDZ domain p
36	56.5	16.3	211	2 T11724	hypothetical prote
37	56.5	16.3	217	2 T27331	hypothetical prote
38	56.5	16.3	256	2 G35866	probable pyrroloqu
39	56.5	16.3	770	2 S00643	anthranilate synth
40	56.5	16.3	884	2 C70729	hypothetical prote
41	56	16.2	651	2 F54457	methionine-tRNA li
42	56	16.2	753	1 D72660	probable aldehyde
43	56	16.2	838	2 T45699	hypothetical prote
44	55.5	16.0	205	2 F87660	hydrolyase, haloact
45	55.5	16.0	299	2 D75138	hypothetical prote

ALIGNMENTS

RESULT 1

AS3019 collagen alpha 1(XVIII) chain - human (fragment)

N/Contains: endostatin

C/Species: Homo sapiens (man)

C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 31-Mar-2000

C/Accession: AS3019

R/Oh: S.P.; Warman, M.L.; Seldin, M.F.; Cheng, S.D.; Knoll, J.H.M.; Timmons, S.; Olsen, Genomics 19, 494-499, 1994

A/Title: Cloning of cDNA and genomic DNA encoding human type XVIII collagen and localiza

A/Reference number: AS3019; MUID:94245237; PMID:8188291

A/Accession: AS3019

A/Molecule type: mRNA

A/Residues: 1-684 <OH>

A/Cross-references: GB:L22548; NID:G348908; PIDN:AA51864.1; PID:G562794

A/Note: The cited accession number, L22548, is not in Genbank release 103

A/Note: In the authors' translation, 482-61y is not shown, residues 483-490 are shifted

C/Comment: Proline and lysines at the third position of the tripeptide repeating unit (

lated and subsequently O-glycosylated.

C/Comment: Different splice forms of collagen alpha 1(XVIII) may be involved in perivasc

C/Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of un

C/Genetics:

A:Gene: GDB:COL18A1

A:Cross-references: GDB:138752; OMIM:120328

A/Map position: 21q22.3-21q22.3

C/Superfamily: unassigned collagens

C/Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteoglyc

F/1-684/Product: collagen alpha 1(XVIII) chain (fragment) #status predicted <MAT>

F/1-59/Domains: collagenous (fragment) #status predicted <CO4>

F/74-115/Domains: collagenous #status predicted <CO5>

F/129-201/Domains: collagenous #status predicted <CO6>

F/212-244/Domains: collagenous #status predicted <CO7>

F/257-278/Domains: collagenous #status predicted <CO8>

F/262-264/Region: cell attachment (R-G-D) motif

F/286-340/Domains: collagenous #status predicted <CO9>

F/354-371/Domains: collagenous #status predicted <CO10>

F/502-684/Product: endostatin #status predicted <BS7>

F/509-684/Region: multiplexin collagen carboxyl-terminal similarity

Query Match 100.0%; Score 346; DB 2; Length 684;

Best Local Similarity 100.0%; Pred. No. 9.9e-34;

Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFSPWALPSGSEGLPKGARRISFGKQVLRHPTWPKSVHSGS 60

DB 565 ADRAAVPIVNLKDELLFSPWALPSGSEGLPKGARRISFGKQVLRHPTWPKSVHSGS 624

QY 61 DPN 63

DB 625 DPN 627

RESULT 2

B56101
 collagen alpha 1(XVII) chain precursor, long splice form - mouse
 M.Comments: collagen alpha 1(XVII) chain precursor, medium splice form; endostatin
 C.Specific: Mus musculus (house mouse)
 C.Date: 03-Oct-1995 #sequence revision 08-May-1998 #text change 31-Mar-2000
 C.Accession: B56101; C56101; S72450; S65595; P06751; A54072; A58816
 R.Rehm, M.; Pihlajaniemi, T.
 J. Biol. Chem. 270, 4705-4711, 1995
 A>Title: Identification of three N-terminal ends of type XVII collagen chains and their
 relation homologous to rat and Drosophila frizzled proteins.
 A.Reference number: A56101; MUID:95181468; PMID:7876242
 A.Accession: B56101
 A.Molecule type: mRNA
 A.Residues: 1-562 <REH1>
 A.Cross-references: GB:U11637, NID:9618429, PIDN:AAC52179.1; PID:9618430
 A.Experimental source: splice form clone PB17.24
 A.Accession: C56101
 A.Molecule type: mRNA
 A.Residues: 1-239, 487-562 <REH2>
 A.Cross-references: GB:U11637, NID:9618429
 A.Experimental source: splice form clones PB8.1, PE19, PE15.2
 R.Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.
 submitted to the EMBL Data Library, August 1993
 A.Reference number: S72450
 A.Accession: S72450
 A.Molecule type: mRNA
 A.Residues: 487-1146, 'L', 1148-1193, 'F', 1195-1210, 'R', 1212-1512, 'L', 1514-1522, 'F', 1524-16
 A.Cross-references: EMBL:L22545; NID:9348968; PIDN:AA19787.1; PID:9511298
 R.Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, A.B.R.
 Proc. Natl. Acad. Sci. U.S.A. 91, 4229-4233, 1994
 A>Title: Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xaa
 A.Reference number: A58370; MUID:94240111; PMID:8183893
 A.Accession: S65595
 A.Molecule type: mRNA
 A.Residues: 487-1146, 'L', 1514-1522, 'F', 1524-1683, 'V', 1685-1774 <OH2>
 A.Cross-references: EMBL:L22545
 R.Abe, N.; Muragaki, Y.; Yoshida, H.; Inoue, H.; Niimura, Y.
 Blochm. Biophys. Res. Commun. 196, 576-582, 1993
 A>Title: Identification of a novel collagen chain represented by extensive interruptions
 A.Reference number: P06751; MUID:94059075; PMID:8240330
 A.Accession: P06751
 A.Molecule type: mRNA
 A.Residues: 635-1774 <ABE>
 R.Rehm, M.; Hincik, E.; Pihlajaniemi, T.
 J. Biol. Chem. 269, 13929-13935, 1994
 A>Title: Primary structure of the alpha1 chain of mouse type XVIII collagen, partial str
 collagen chain.
 A.Reference number: A54072; MUID:94245707; PMID:8186673
 A.Accession: A54072
 A.Molecule type: DNA; mRNA
 A.Residues: 1293-1403, 'R', 1405-1774 <REH3>
 A.Cross-references: GB:U03714; NID:9487733; PIDN:AA20657.1; PID:9487734
 R.O'Reilly, M.S.; Boehm, T.; Shing, Y.; Fukui, N.; Vasios, G.; Lane, W.S.; Flynn, E.; B
 Cell 88, 277-285, 1997
 A>Title: Endostatin, an endogenous inhibitor of angiogenesis and tumor growth.
 A.Reference number: A58816; MUID:97160848; PMID:9008168
 A.Accession: A58816
 A.Molecule type: protein
 A.Residues: 1591-1610 <ORE>
 A.Experimental source: hemangioendothelium cells
 A>Note: Inhibits endothelial cell proliferation
 C.Comment: Prolines and tyrosines at the third position of the tripeptide repeating unit
 lated and subsequently O-glycosylated.
 C.Comment: The different splice forms of collagen alpha 1(XVII) may be involved in per
 C.Comment: Endostatin is released from collagen alpha 1(XVII) chain by the action of un
 ay be useful in treating solid tumors.
 C.Genetics:
 A.Gene: MGI:Coll18a1
 A.Cross-references: MGI:71175
 A.Map position: 10:61.0
 A.Introns: 1295/3; 1310/1; 1331/1; 1345/3; 1388/3; 1437/1; 1461/3; 1505/3; 1516/3; 1599/

A>Note: the list of introns is incomplete

C.Superfamily: unassigned collagens

C.Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteogly

F1-1774/Product: collagen alpha 1(XVII) chain precursor, long splice form #status pred

F1-239, 487-1774/Product: collagen alpha 1(XVII) chain precursor, medium splice form #

F1-24/Domin: signal sequence #status predicted <SIG>

F1-486/Region: frizzled similarity

F1-786-812/Domin: collagenous #status predicted <CO01>

F1-823-896/Domin: collagenous #status predicted <CO02>

F1-921-1042/Domin: collagenous #status predicted <CO03>

F1-1066-1148/Domin: collagenous #status predicted <CO04>

F1-1163-1204/Domin: collagenous #status predicted <CO05>

F1-1218-1290/Domin: collagenous #status predicted <CO06>

F1-1301-1333/Domin: collagenous #status predicted <CO07>

F1-1346-1369/Domin: collagenous #status predicted <CO08>

F1-1351-1353/Region: cell attachment (R-G-D) motif

F1-1377-1428/Domin: collagenous #status predicted <CO09>

F1-1442-1459/Domin: collagenous #status predicted <CO10>

F1-1591-1774/Product: endostatin #status predicted <EST>

F1-1598-1774/Region: multiplexin collagen carboxyl-terminal similarity

F1-354, 361, 947/Binding site: carbohydrate (Asn) (covalent) #status predicted

F1-699, 704, 1716/Binding site: carbohydrate (Ser) (covalent) #status predicted

F1-910, 913, 1053/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 85.0% Score 294; DB 2; Length 1774;

Best Local Similarity 81.0% Pred. No. 5, 9e-27;

Matches 51; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ADRAAVPIVNLKDELLFSGSECPKPGARIFSPOKDVLRHPTMPQKSVHGS 60

Db 1654 ADRGSPVIVNLKDEVLSPMSLFSGSGQGLPGARIFSPOKDVLRHPTMPQKSVHGS 1713

Qy 61 DPN 63

Db 1714 DPS 1716

RESULT 3

B56101
 collagen alpha 1(XVIII) chain precursor, short splice form - mouse
 N.Comments: endostatin
 C.Specific: Mus musculus (house mouse)
 C.Date: 03-Oct-1995 #sequence revision 08-May-1998 #text change 31-Mar-2000
 C.Accession: A56101; A58371; S72450; S65595
 R.Rehm, M.; Pihlajaniemi, T.
 J. Biol. Chem. 270, 4705-4711, 1995
 A>Title: Identification of three N-terminal ends of type XVIII collagen chains and their
 relation homologous to rat and Drosophila frizzled proteins.
 A.Reference number: A56101; MUID:94240112; PMID:8183894
 A.Accession: A56101
 A.Molecule type: mRNA
 A.Residues: 1-103 <REH1>
 A.Cross-references: GB:U11636; NID:9618427; PIDN:AAC52178.1; PID:9618428
 R.Rehm, M.; Pihlajaniemi, T.
 Proc. Natl. Acad. Sci. U.S.A. 91, 4234-4238, 1994
 A>Title: Alpha1(XVIII), a collagen chain with frequent interruptions in the collagenous
 A.Reference number: A58371; MUID:94240112; PMID:8183894
 A.Accession: A58371
 A.Molecule type: mRNA
 A.Residues: 1-928 <REH2>
 A.Cross-references: GB:L16898; NID:9404754; PIDN:AA37434.1; PID:9553894
 R.Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.
 submitted to the EMBL Data Library, August 1993
 A.Reference number: S72450
 A.Accession: S72450
 A.Molecule type: mRNA
 A.Residues: 28-687, 'L', 689-734, 'F', 736-751, 'R', 753-1315 <OHM>
 A.Cross-references: EMBL:L22545; NID:9348968; PIDN:AA19787.1; PID:9511298
 R.Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, A.B.R.
 Proc. Natl. Acad. Sci. U.S.A. 91, 4229-4233, 1994
 A>Title: Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xaa
 A.Reference number: A58370; MUID:94240111; PMID:8183893
 A.Accession: S65595

[illegible]

```

09      24   FSGSEGFLKCGARIFSFQKDVLK---HPTWPKQSVMHGSD    61
          ||| |||| |:-: |||||
Db       13 FAGKAVALSPGASVEFGTGRVTKRYKVPFGPSGPMYGS    53

RESULT 9
T17408
rRNA (adenine-N6-)-methyltransferase (EC 2.1.1.48) 1 - Streptomyces venezuelae
C:Species: Streptomyces venezuelae
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 08-Sep-2000
C:Accession: T17408
C:Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
A>Title: A gene cluster for macrolide antibiotic biosynthesis in streptomycetes venezuelae
A:Reference number: Z18773; MUID:9844533; PMID:9770448
A:Accession: T17408
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-336 <XUE>
A:Cross-references: EMBL:AF079138; NID:g3808326; PID:g3800833; PIDN:AAC69328.1
C:Genetics:
A:Gene: pikR1
C:Keywords: methyltransferase; S-adenosylmethione

Query Match      18.2%; Score 63; DB 2; Length 336;
Best Local Similarity 40.0%; Pred. No. 7.1;
Matches 14; Conservative 3; Mismatches 18; Indels 0; Gaps 0;

09      1 ADRAAPIVNLKDQLFPSSWEALFSGSEGLKKGCA 35
          ||| |:-: |||||
Db     235 ADRHGTFPDGGLPKDUTLTOWIALFOASQSPSYAECA 269

```

```

RESULT 10
T26705
hypothetical protein Y38H6C.17 - Caenorhabditis elegans
C|Species: Caenorhabditis elegans
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 12-May-2003
C|Accession: T26705
R|White, S.
submitted to the EMBL Data Library, September 1998
A|Reference number: Z20255
A|Accession: T26705
A|Status: preliminary; translated from GB/EMBL/DDBU
A|Molecule type: DNA
A|Residues: 1-467 <WIL>
A|Cross-references: EMBL:AL031630; PIDN:CAA0995.1; GSPDB:IGN00023; CESP:Y38H6C.17
A|Experimental source: clone Y38H6C
C|Genetics:
A|Gene: CESP:Y38H6C.17
A|Map position: 5
A|Introns: 9/2; 49/2; 79/3; 109/2; 140/2; 168/3; 225/3; 322/2; 382/1; 429/2
C|Superfamily: Arabidopsis amino acid transport protein I

Query Match      18.1% Score 62.5; DB 2; Length 467;
Best Local Similarity 33.9%; Pred. NO. 12;
Matches 20; Conservative 12; Mismatches 12; Indels 15; Gaps 5;

QY      4 AAVPIVVLKKELT----FPGME-ALFGSGEG-PLKPGARIFSPDG-----KDVLRRHP 49
          |||::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db       218 AAIALIT--DELLSHTWPTWEIDSTIGVEGVSILAASGLYSFEGQAMVPLENSLRNP 274

RESULT 11
FE9797
conserved hypothetical protein yeast - Bacillus subtilis
C|Species: Bacillus subtilis
C|Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C|Accession: FE9797
R|Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azavedo, V.; Berteaux, D.; Bron, S.; Broutier, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chom, A.; Ehrlich, S.D.; Emmertson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
```

Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
 Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidis, A.; Lardinois
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mause
 Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akouchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A65580; MUID:98044033; PMID:9384377
 A:Accession: F69797
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-620 <KUN>
 A:Cross-references: GB:Z99107; GB:AL009126; NID:q2632866; PIDN:CAB12524.1; PID:g2633018
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yesw

Query Match 18.1%; Score 62.5; DB 2; Length 620;
 Best Local Similarity 23.5%; Pred. No. 17;
 Matches 23; Conservative 10; Mismatches 16; Indels 49; Gaps 4;

QY 12 KDELPPSW-----EALFSGSEGLPKFGARIFS----- 39

DB 375 KQELIRGMAVVDHDKGMYSTGLGHGDLADHTGDLDRGRGLVFFVGHEDKNAKYGLSPFD 434

QY 40 -----FDGKDYLR-----HPTWPKSYW-HGS 60
 DB 435 AATGKLWGVYAGKDVGRGMAADIDPRYPQGEWANGS 472

RESULT 12

hypothetical protein C24A8.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T25565

R:Nelson, J.; Gatlung, S.

submitted to the EMBL Data Library, December 1996

A:Description: The sequence of *C. elegans* cosmid C24A8.

A:Reference number: Z20051

A:Accession: T25565

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-419 <NEI>

A:Cross-references: EMBL:U80845; PIDN:AB37978.1; GSPDB:GN00028; CESP:C24A8.1

A:Experimental source: strain Bristol N2; clone C24A8

C:Genetics:

A:Gene: CESP:C24A8.1

A:Map position: X

A:Interon: 27/3; 49/3; 87/2; 186/3; 213/1; 250/3; 283/1; 348/3; 374/2; 407/1

C:Superfamily: vertebrate rhodopsin

Query Match 17.9%; Score 62; DB 2; Length 419;
 Best Local Similarity 43.2%; Pred. No. 12;
 Matches 16; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 4 AAAPVNLKDELLFPGSEALFSGSEGLPKFGARIFS 40
 DB 8 AAAPVNLKDELLFPGSEALFSGSEGLPKFGARIFS 44

RESULT 13

serine proteinase (EC 3.4.21.-) precursor - *Serratia marcescens* (strain IFO-3046)

C:Species: *Serratia marcescens*
 C:Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 22-Jun-1999

C:Accession: A29840
 R:Yanagida, N.; Uozumi, T.; Beppu, T.

J. Bacteriol. 166, 937-944, 1986
 A>Title: Specific excretion of *Serratia marcescens* protease through the outer membrane o
 A:Reference number: A29840; MUID:86223815; PMID:3011754
 A:Accession: A29840
 A:Molecule type: DNA
 A:Residues: 1-1045 <YAN>
 A:Cross-references: GB:M13469; NID:9152857; PIDN:AAA26572.1; PID:9152858
 C:Superfamily: *Serratia* serine proteinase; subtilisin homology
 C:Keywords: hydrolase; serine proteinase
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-408/Product: serine proteinase #status predicted <MAT>
 F:67-355/Domain: subtilisin homology <SBT>
 F:409-1045/Domain: carboxyl-terminal propeptide #status predicted <PRO>
 F:76,112,341/Active site: Asp, His, Ser #status predicted

Query Match 17.6%; Score 61; DB 2; Length 1045;
 Best Local Similarity 34.9%; Pred. No. 48;
 Matches 15; Conservative 4; Mismatches 24; Indels 0; Gaps 0;

QY 14 ELLFPGSEALFSGSEGLPKFGARIFSFDGKDYLRHPTWPKSYW 56
 DB 449 ECSFDSNGNDISGHGGLTKTGAGTLALGNNTYRSDTWKGV 491

RESULT 14

probable transketolase (EC 2.2.1.1) [imported] - *Sinorhizobium meliloti* (strain 1021) ma

C:Species: *Sinorhizobium meliloti*

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001

C:Accession: A95866

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A>Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11491431

A:Accession: A95866

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-694 <KUR>

A:Cross-references: GB:AL591985; PIDN:CAC48593.1; PID:91514065; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid pSymB

R:Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.,

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kallman, S.; Keating, D.H.; Kiss, E.; Komp, C.; LeLaure,

hebaunt, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K

A>Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: cbhT; SMB20200

A:Genome: plasmid

C:Superfamily: transketolase; thiamin pyrophosphate-binding domain homology

C:Keywords: transketase

Query Match 17.3%; Score 60; DB 2; Length 694;
 Best Local Similarity 29.2%; Pred. No. 39;
 Matches 21; Conservative 8; Mismatches 27; Indels 16; Gaps 3;

QY 3 RAAPVNLKDELLFPGSEALFSGSEGLPKFGARIFSFDGKDYLR-----RHPTWPK 53
 DB 10 RAAPVNLKDELLFPGSEALFSGSEGLPKFGARIFSFDGKDYLR-----RHPTWPK 67

QY 54 K-----SYMHGS 60
 DB 68 KDFVLSKAGHS 79

RESULT 15

T4496

cellulose 1,4-beta-cellulohydrolase (EC 3.2.1.91) precursor [imported] - *Thermomonospora f*

N:Alternate names: beta-1,4-exocellulase B6 precursor

C:Species: *Thermomonospora fusca*

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using BW model

Run on: September 10, 2003, 14:10:33 ; Search time 26 Seconds

(without alignments)
353.557 Million cell updates/sec

Title: US-09-822-540A-1

Perfect score: 346
Sequence: 1 ADRAAVPIVNLKDELLFPPSW.....DVLRRPTWPKSVHMGSDPN 63

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTIS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	346	100.0	63	9 US-09-822-540A-1	Sequence 1, Appli
2	346	100.0	178	15 US-10-131-241-60	Sequence 60, Appli
3	346	100.0	178	15 US-10-042-347-5	Sequence 5, Appli
4	346	100.0	179	15 US-10-131-241-57	Sequence 57, Appli
5	346	100.0	180	15 US-10-131-241-56	Sequence 56, Appli
6	346	100.0	181	15 US-10-131-241-55	Sequence 55, Appli
7	346	100.0	182	10 US-09-998-831-14	Sequence 14, Appli
8	346	100.0	182	15 US-10-131-241-54	Sequence 54, Appli
9	346	100.0	182	15 US-10-042-347-3	Sequence 3, Appli
10	346	100.0	183	9 US-09-873-676-2	Sequence 2, Appli
11	346	100.0	183	12 US-10-292-418-4	Sequence 4, Appli
12	346	100.0	183	14 US-10-080-797-1	Sequence 1, Appli
13	346	100.0	183	15 US-10-131-241-52	Sequence 52, Appli
14	346	100.0	684	11 US-09-961-403-5	Sequence 5, Appli
15	346	100.0	1516	15 US-10-060-036-166	Sequence 166, Appli

16	338	97.7	180	15 US-10-131-241-47	Sequence 47, Appli
17	302	87.3	184	12 US-10-292-418-35	Sequence 35, Appli
18	302	87.3	184	12 US-09-938-391-4	Sequence 4, Appli
19	302	87.3	184	15 US-10-131-241-49	Sequence 49, Appli
20	302	87.3	230	12 US-09-938-391-2	Sequence 2, Appli
21	294	85.0	191	10 US-09-998-831-13	Sequence 13, Appli
22	294	85.0	207	14 US-10-080-797-3	Sequence 3, Appli
23	291	84.1	184	15 US-10-292-418-18	Sequence 18, Appli
24	291	84.1	184	15 US-10-131-241-46	Sequence 46, Appli
25	277.5	80.2	185	15 US-10-036-869-36	Sequence 36, Appli
26	182	52.6	31	9 US-09-822-540A-2	Sequence 2, Appli
27	94	27.2	16	10 US-09-766-412-32	Sequence 32, Appli
28	67.5	19.5	614	15 US-10-146-473-45	Sequence 45, Appli
29	63	18.2	336	10 US-09-861-289-27	Sequence 27, Appli
30	63	18.2	336	10 US-09-860-846-27	Sequence 27, Appli
31	63	18.2	336	11 US-09-988-384B-27	Sequence 27, Appli
32	63	18.2	336	11 US-09-836-821-27	Sequence 27, Appli
33	63	18.2	11877	10 US-09-861-289-6	Sequence 6, Appli
34	63	18.2	11877	10 US-09-860-846-6	Sequence 6, Appli
35	63	18.2	11877	11 US-09-836-821-6	Sequence 6, Appli
36	63	18.2	12199	11 US-09-988-384B-6	Sequence 6, Appli
37	62.5	18.1	620	15 US-10-124-880-6	Sequence 6, Appli
38	61.5	17.8	3170	15 US-10-128-714-8504	Sequence 8504, Ap
39	59	17.1	14	10 US-09-766-412-31	Sequence 31, Appli
40	57.5	16.6	506	12 US-09-882-227-116	Sequence 116, App
41	57.5	16.6	553	11 US-09-292-862-2	Sequence 2, Appli
42	57	16.5	636	15 US-10-156-761-13432	Sequence 13432, A
43	57	16.5	2037	10 US-09-951-401-3	Sequence 3, Appli
44	57	16.5	2037	10 US-09-922-101-3	Sequence 3, Appli
45	57	16.5	2037	10 US-09-951-402-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-09-822-540A-1
; Sequence 1, Application US/09822540A
; Patent No. US20020058620A1
; GENERAL INFORMATION:
; APPLICANT: Phoenix Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cell Growth Regulation System
; FILE REFERENCE: PhoenixPharFull1
; CURRENT APPLICATION NUMBER: US/09/822,540A
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/194,561
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-540A-1

Query Match      100.0%; Score 346; DB 9; Length 63;
Best Local Similarity 100.0%; Pred. No. 2.1e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ADRAAVPIVNLKDELLFPPSWALFSGSEGPLKPGARIFSPDGKXVLRHPTWPKSVHMG 60
      |||
Db      1 ADRAAVPIVNLKDELLFPPSWALFSGSEGPLKPGARIFSPDGKXVLRHPTWPKSVHMG 60

Qy      61 DPN 63
      |||
Db      61 DPN 63

RESULT 2
US-10-131-241-60
; Sequence 60, Application US/10131241
; Publication No. US20030012792A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-60

Query Match          100.0%; Score 346; DB 15; Length 178;
Best Local Similarity 100.0%; Pred. No. 6.5e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAIVPNLKDLELFPSEWALFSGSEGLPKGARIFSPDGKDLRHPTWPKSVHGS 60
DB 60 ADRAAIVPNLKDLELFPSEWALFSGSEGLPKGARIFSPDGKDLRHPTWPKSVHGS 119

QY 61 DPN 63
DB 120 DPN 122

RESULT 3
US-10-042-347-5
; Sequence 5, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide Fr
; FILE REFERENCE: 05213-0880 (43170-249874)
; CURRENT APPLICATION NUMBER: US/10/042,347
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-347-5

Query Match          100.0%; Score 346; DB 15; Length 178;
Best Local Similarity 100.0%; Pred. No. 6.5e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 ADRAAIVPNLKDLELFPSEWALFSGSEGLPKGARIFSPDGKDLRHPTWPKSVHGS 60
DB 60 ADRAAIVPNLKDLELFPSEWALFSGSEGLPKGARIFSPDGKDLRHPTWPKSVHGS 119

QY 61 DPN 63
DB 120 DPN 122

RESULT 4
US-10-131-241-57
; Sequence 57, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Fortier, Anne H.
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-57

Query Match          100.0%; Score 346; DB 15; Length 179;
Best Local Similarity 100.0%; Pred. No. 6.5e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAIVPNLKDLELFPSEWALFSGSEGLPKGARIFSPDGKDLRHPTWPKSVHGS 60
DB 60 ADRAAIVPNLKDLELFPSEWALFSGSEGLPKGARIFSPDGKDLRHPTWPKSVHGS 119

QY 61 DPN 63
DB 120 DPN 122

RESULT 5
US-10-131-241-56
; Sequence 56, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-56
```

```
Query Match          100.0%; Score 346; DB 15; Length 180;
Best Local Similarity 100.0%; Pred. No. 6.6e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPGSEGLPKGARIFFDGKDVLRHPTWPKSVHGS 60
DB 64 ADRAAVPIVNLKDELLFPGSEGLPKGARIFFDGKDVLRHPTWPKSVHGS 123

QY 61 DPN 63
DB 124 DPN 126

RESULT 6
US-10-131-241-55
; Sequence 55, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent in version 3.1
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-55

Query Match          100.0%; Score 346; DB 15; Length 181;
Best Local Similarity 100.0%; Pred. No. 6.6e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPGSEGLPKGARIFFDGKDVLRHPTWPKSVHGS 60
DB 64 ADRAAVPIVNLKDELLFPGSEGLPKGARIFFDGKDVLRHPTWPKSVHGS 123

QY 61 DPN 63
DB 124 DPN 126

RESULT 7
US-09-998-831-14
; Sequence 14, Application US/09998831
; Patent No. US20020119153A1
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/998,831
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 09/561,108
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 14
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-998-831-14

Query Match          100.0%; Score 346; DB 10; Length 182;
Best Local Similarity 100.0%; Pred. No. 6.7e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPGSEGLPKGARIFFDGKDVLRHPTWPKSVHGS 60
DB 64 ADRAAVPIVNLKDELLFPGSEGLPKGARIFFDGKDVLRHPTWPKSVHGS 123

QY 61 DPN 63
DB 124 DPN 126

RESULT 8
US-10-131-241-54
; Sequence 54, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 54
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-54

Query Match          100.0%; Score 346; DB 15; Length 182;
Best Local Similarity 100.0%; Pred. No. 6.7e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPGSEGLPKGARIFFDGKDVLRHPTWPKSVHGS 60
DB 64 ADRAAVPIVNLKDELLFPGSEGLPKGARIFFDGKDVLRHPTWPKSVHGS 123

QY 61 DPN 63
DB 124 DPN 126

RESULT 9
US-10-042-347-3
; Sequence 3, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide F
; FILE REFERENCE: 05213-0880 (43170-249874)
; CURRENT APPLICATION NUMBER: US/10/042,347
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
```

```

; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-347-3
```

```

Query Match          100.0%; Score 346; DB 15; Length 182;
Best Local Similarity 100.0%; Pred. No. 6,7e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSFQDKDYLRRPTWPKSVWHS 60
DB 64 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSFQDKDYLRRPTWPKSVWHS 123
QY 61 DPN 63
DB 124 DPN 126
```

```

RESULT 10
US-09-873-676-2
; Sequence 2, Application US/09873676
; Patent No. US2002077289A1
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-25333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-873-676-2
```

```

Query Match          100.0%; Score 346; DB 9; Length 183;
Best Local Similarity 100.0%; Pred. No. 6,7e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSFQDKDYLRRPTWPKSVWHS 60
DB 64 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSFQDKDYLRRPTWPKSVWHS 123
QY 61 DPN 63
DB 124 DPN 126
```

```

RESULT 11
US-10-292-418-4
; Sequence 4, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
```

```

; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-418-4
```

```

Query Match          100.0%; Score 346; DB 12; Length 183;
Best Local Similarity 100.0%; Pred. No. 6,7e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSFQDKDYLRRPTWPKSVWHS 60
DB 64 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSFQDKDYLRRPTWPKSVWHS 123
QY 61 DPN 63
DB 124 DPN 126
```

```

RESULT 12
US-10-080-797-1
; Sequence 1, Application US/10080797
; Publication No. US20020183253A1
; GENERAL INFORMATION:
; APPLICANT: Campochiaro, Peter A.
; APPLICANT: Dixon, Katherine H.
; APPLICANT: Brazzell, Romulus K.
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR
; FILE REFERENCE: 4-31881A
; CURRENT APPLICATION NUMBER: US/10/080,797
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Human
US-10-080-797-1
```

```

Query Match          100.0%; Score 346; DB 14; Length 183;
Best Local Similarity 100.0%; Pred. No. 6,7e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSFQDKDYLRRPTWPKSVWHS 60
DB 64 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSFQDKDYLRRPTWPKSVWHS 123
QY 61 DPN 63
DB 124 DPN 126
```

```

RESULT 13
US-10-131-241-52
; Sequence 52, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
```

```

; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-52
```

```

Query Match          100.0%; Score 346; DB 15; Length 183;
Best Local Similarity 100.0%; Pred. No. 6.7e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy      1 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHS 60
Db      64 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHS 123
```

```

Qy      61 DPN 63
Db      124 DPN 126
```

```

RESULT 14
US-09-961-403-5
; Sequence 5, Application US/09961403
; Publication No. US2003007589A1
; GENERAL INFORMATION:
; APPLICANT: HE-STUMP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PRDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-5
```

```

Query Match          100.0%; Score 346; DB 11; Length 684;
Best Local Similarity 100.0%; Pred. No. 2.9e-34;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy      1 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHS 60
Db      565 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHS 624
```

```

Qy      61 DPN 63
Db      625 DPN 627
```

```

RESULT 15
US-10-060-036-166
; Sequence 166, Application US/10060036
; Publication No. US20030073144A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 166
; LENGTH: 1516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-036-166
```

```

Query Match          100.0%; Score 346; DB 15; Length 1516;
Best Local Similarity 100.0%; Pred. No. 7e-34;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy      1 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHS 60
Db      1397 ADRAAVPIVNLKDELLFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHS 1456
```

```

Qy      61 DPN 63
Db      1457 DPN 1459
```

```

Search completed: September 10, 2003, 14:13:44
Job time : 27 secs
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2003, 14:09:53; Search time 21 Seconds
(without alignments)
(126.933 Million cell updates/sec)

Title: US-09-822-540A-1

Perfect score: 346

Sequence: 1 ADRAAVPIVNLKDELLFSPSW.....DVLRRHPWQKSWHGDSPN 63

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUTS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	346	100.0	178	4	US-09-315-689-5
2	346	100.0	182	4	US-09-561-500-14
3	346	100.0	182	4	US-09-561-108-14
4	346	100.0	182	4	US-09-315-689-3
5	346	100.0	182	4	US-09-561-526-14
6	346	100.0	182	4	US-09-561-499-14
7	346	100.0	183	3	US-09-206-059-2
8	294	85.0	191	4	US-09-561-500-13
9	294	85.0	191	4	US-09-561-108-13
10	294	85.0	191	4	US-09-561-526-13
11	294	85.0	191	4	US-09-561-499-13
12	291	84.1	195	1	US-08-159-784-2
13	277.5	80.2	185	3	US-08-985-526-36
14	190	54.9	191	1	US-08-159-784-3
15	94	27.2	16	3	US-09-385-442-32
16	63	18.2	336	3	US-09-105-537-27
17	63	18.2	11877	3	US-09-105-537-6
18	62.5	18.1	620	4	US-09-311-626B-6
19	59.5	17.2	92	4	US-09-673-809-106
20	59	17.1	14	3	US-09-385-442-31
21	59	17.1	450	4	US-09-252-991A-18659
22	59	17.1	1044	4	US-09-552-991A-18853
23	57.5	16.6	506	2	US-08-849-480A-5
24	57.5	16.6	553	3	US-09-083-351-2
25	57.5	16.6	553	3	US-09-083-352-2
26	57	16.0	2037	3	US-09-306-998-3
27	55.5	16.0	191	4	US-09-376-728-2

28	55	15.9	444	1	US-08-483-140-28	Sequence 28, Appl
29	55	15.9	444	2	US-08-485-938A-32	Sequence 32, Appl
30	54.5	15.8	470	2	US-08-377-440A-1	Sequence 1, Appl
31	54.5	15.8	470	3	US-08-440-530-1	Sequence 1, Appl
32	54.5	15.8	578	1	US-08-653-740-3	Sequence 3, Appl
33	54.5	15.8	578	2	US-09-073-594-3	Sequence 3, Appl
34	54.5	15.8	578	3	US-09-275-925-3	Sequence 3, Appl
35	54.5	15.8	636	1	US-08-653-740-5	Sequence 3, Appl
36	54.5	15.8	636	2	US-09-073-594-5	Sequence 5, Appl
37	54.5	15.8	636	3	US-09-275-925-5	Sequence 5, Appl
38	54	15.6	392	4	US-09-996-243-205	Sequence 205, App
39	54	15.6	393	4	US-09-482-273-154	Sequence 154, App
40	54	15.6	878	4	US-09-463-238-4	Sequence 4, Appl
41	53.5	15.5	883	4	US-09-463-238-19	Sequence 19, Appl
42	53.5	15.5	396	4	US-09-198-452A-147	Sequence 147, App
43	53.5	15.5	632	4	US-09-252-991A-26091	Sequence 26091, A
44	53	15.3	312	3	US-09-216-295-21	Sequence 21, Appl
45	53	15.3	371	3	US-09-104-308-1	Sequence 1, Appl

ALIGNMENTS

```

RESULT 1
US-09-315-689-5
; Sequence 5, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-689-5

Query Match      100.0%; Score 346; DB 4; Length 178;
Best Local Similarity 100.0%; Pred. No. 3e-38;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ADRAAVPIVNLKDELLFSPSWALFSGSEGLPKGARIFSDGKDVLRHPTWQKSWHGS 60
      |||||||
DB      60 ADRAAVPIVNLKDELLFSPSWALFSGSEGLPKGARIFSDGKDVLRHPTWQKSWHGS 119

QY      61 DPN 63
      |||
DB      120 DPN 122

RESULT 2
US-09-561-500-14
; Sequence 14, Application US/09561500
; Patent No. 6342219
; GENERAL INFORMATION:
; APPLICANT: Phillip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002500
; CURRENT APPLICATION NUMBER: US/09/561,500
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 182
; TYPE: PRT

```

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
OTHER INFORMATION: PEPTIDE
US-09-561-500-14

Query Match 100.0%; Score 346; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.1e-38;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHS 60
DB 64 ADRAAVPIVNLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHS 123

QY 61 DPN 63
DB 124 DPN 126

RESULT 3
US-09-561-108-14
Sequence 14, Application US/09561108
Patent No. 6342221
GENERAL INFORMATION:
APPLICANT: Philip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002584
CURRENT APPLICATION NUMBER: US/09/561,108
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIORITY FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 182
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
OTHER INFORMATION: PEPTIDE
US-09-561-108-14

Query Match 100.0%; Score 346; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.1e-38;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHS 60
DB 64 ADRAAVPIVNLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHS 123

QY 61 DPN 63
DB 124 DPN 126

RESULT 4
US-09-315-689-3
Sequence 3, Application US/09315689
Patent No. 6346510
GENERAL INFORMATION:
APPLICANT: Folkman, Judah
APPLICANT: O'Reilly, Michael
TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
FILE REFERENCE: 05213-0229
CURRENT APPLICATION NUMBER: US/09/315,689
PRIOR FILING DATE: 1999-05-20
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 182
TYPE: PRT
ORGANISM: Homo sapiens

US-09-315-689-3

Query Match 100.0%; Score 346; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.1e-38;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHS 60
DB 64 ADRAAVPIVNLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHS 123

QY 61 DPN 63
DB 124 DPN 126

RESULT 5
US-09-561-526-14
Sequence 14, Application US/09561526
Patent No. 6416758
GENERAL INFORMATION:
APPLICANT: Philip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002586
CURRENT APPLICATION NUMBER: US/09/561,526
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIORITY FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 182
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
OTHER INFORMATION: PEPTIDE
US-09-561-526-14

Query Match 100.0%; Score 346; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.1e-38;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHS 60
DB 64 ADRAAVPIVNLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHS 123

QY 61 DPN 63
DB 124 DPN 126

RESULT 6
US-09-561-499-14
Sequence 14, Application US/09561499
Patent No. 6524583
GENERAL INFORMATION:
APPLICANT: Philip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002582
CURRENT APPLICATION NUMBER: US/09/561,499
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIORITY FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 182
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC

OTHER INFORMATION: PEPTIDE
US-09-561-499-14

Query Match 100.0%; Score 346; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 3,1e-38;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPGSEGLPKPGARIFSPDGKDVLRHPTWPKSVMHGS 60
|||
Db 64 ADRAAVPIVNLKDELLFPGSEGLPKPGARIFSPDGKDVLRHPTWPKSVMHGS 123

QY 61 DPN 63
|||
Db 124 DPN 126

RESULT 7
US-09-206-059-2
Sequence 2, Application US/09206059
Patent No. 6201104

GENERAL INFORMATION:
APPLICANT: Macdonald, Nicholas
APPLICANT: Sim, Kim Lee
TITLE OF INVENTION: Angiogenesis-Inhibiting Protein Binding Peptides and
FILE REFERENCE: 05213-0370
CURRENT APPLICATION NUMBER: US/09/206,059
CURRENT FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 80
SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 2
LENGTH: 183
TYPE: PRT
ORGANISM: Homo sapiens
US-09-206-059-2

Query Match 100.0%; Score 346; DB 3; Length 183;
Best Local Similarity 100.0%; Pred. No. 3,1e-38;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPGSEGLPKPGARIFSPDGKDVLRHPTWPKSVMHGS 60
|||
Db 64 ADRAAVPIVNLKDELLFPGSEGLPKPGARIFSPDGKDVLRHPTWPKSVMHGS 123

QY 61 DPN 63
|||
Db 124 DPN 126

RESULT 8
US-09-561-500-13
Sequence 13, Application US/09561500
Patent No. 6342219

GENERAL INFORMATION:
APPLICANT: Philip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002500
CURRENT APPLICATION NUMBER: US/09/561,500
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIOR FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 13
LENGTH: 191
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-500-13

Query Match 85.0%; Score 294; DB 4; Length 191;
Best Local Similarity 81.0%; Pred. No. 2,6e-31;
Matches 51; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Best Local Similarity 81.0%; Pred. No. 2,6e-31;
Matches 51; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPGSEGLPKPGARIFSPDGKDVLRHPTWPKSVMHGS 60
|||
Db 71 ADRAAVPIVNLKDELLFPGSEGLPKPGARIFSPDGKDVLRHPTWPKSVMHGS 130

QY 61 DPN 63
|||
Db 131 DPS 133

RESULT 9
US-09-561-108-13
Sequence 13, Application US/09561108
Patent No. 6342221

GENERAL INFORMATION:
APPLICANT: Philip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002584
CURRENT APPLICATION NUMBER: US/09/561,108
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIOR FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 13
LENGTH: 191
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-108-13

Query Match 85.0%; Score 294; DB 4; Length 191;
Best Local Similarity 81.0%; Pred. No. 2,6e-31;
Matches 51; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPGSEGLPKPGARIFSPDGKDVLRHPTWPKSVMHGS 60
|||
Db 71 ADRAAVPIVNLKDELLFPGSEGLPKPGARIFSPDGKDVLRHPTWPKSVMHGS 130

QY 61 DPN 63
|||
Db 131 DPS 133

RESULT 10
US-09-561-526-13
Sequence 13, Application US/09561526
Patent No. 6416758

GENERAL INFORMATION:
APPLICANT: Philip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002586
CURRENT APPLICATION NUMBER: US/09/561,526
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIOR FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 13
LENGTH: 191
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-526-13

Query Match 85.0%; Score 294; DB 4; Length 191;
Best Local Similarity 81.0%; Pred. No. 2,6e-31;

Query Match	80.2%;	Score 277.5;	DB 3;	Length 185
Best Local Similarity	76.6%;	Pred. No. 3.8e-29;		

Matches	49; Conservative	10; Mismatches	4; Indels	1; Gaps	1;							
QY	1	ADRAAVTIV	NLKDELLF	PSWEALF	SGSGEGGLK	GCARIF	SFDGKDLV	LHP	TPMOKS	VWNG	59	
Db	64	ADRSVIVONT	LRDEVLV	SPSWDSL	FSGSGG	OGLOLGC	ARIF	SFDGKDLV	LHP	PMPORS	VWNG	123
QY	60	SDPN	63									
Db	124	SDPS	127									

RESULT 14
US-08-159

```

1 Sequence 3, Application US/08159784
2 Patent No. 5643783
3
4 GENERAL INFORMATION:
5 APPLICANT: Bjorn R. Olsen
6 TITLE OF INVENTION: NOVEL COLLAGEN AND U
7 NUMBER OF SEQUENCES: 9
8 CORRESPONDENCE ADDRESS:
9 ADDRESSEE: Fish & Richardson
10 STREET: 225 Franklin Street
11 CITY: Boston
12 STATE: Massachusetts
13 COUNTRY: U.S.A.
14 ZIP: 02110-2804
15
16 COMPUTER READABLE FORM:
17 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
18 COMPUTER: IBM PS/2 Model 502 or 55SX
19 OPERATING SYSTEM: MS-DOS (Version 5.0)
20 SOFTWARE: Wordperfect (Version 5.1)
21
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/08/159,784
24 FILING DATE: December 1, 1993
25 CLASSIFICATION: 530
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER:
28 FILING DATE:
29
30 ATTORNEY/AGENT INFORMATION:
31 NAME: John F. Freeman
32 REGISTRATION NUMBER: 29,066
33 REFERENCE/DOCKET NUMBER: 00246/170001
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: (617) 542-5070
36 TELEFAX: (617) 542-8906
37
38 TELEX: 200154
39
40 INFORMATION FOR SEQ ID NO: 3:
41 SEQUENCE CHARACTERISTICS:
42 LENGTH: 191
43 TYPE: amino acid
44 STRADEDNESS: N/A
45 TOPOLOGY: N/A
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
```

Query_Match	54.9%	Score 130;	DB 1;	Length 191;
Best Local Similarity	49.2%;	Pred. No. 1,6e-17;		
Match 31; Conservative	16;	Mismatches 16;	Indels 0;	Gaps 0;

Oy 1 ADEAAVPIVLKDELFPSPWEALFSGSEGPLKPGARIFSFDQDVLRHPTWPQKSVNHGS 60
|-::|::|||::|::|::|||::|::|::|||::|::|::|||::|::|::|||::|::|::|||
71 AEHYSLPILVKGOVLFNWWDISFGSHGGCGENMHLPYISFCGRDINTDPSPDQKVIWHGS 130

QY	61	DPN	63
		:	
Db	131	SPH	133

RESULT 15
US-09-385-442-32
; Sequence 32, Application US/09385442
; Patent No. 6200954
; GENERAL INFORMATION:
; APPLICANT: Ge, Ruowen

```

1. APPLICANT: Kini, R. Manjunatha
2. TITLE OF INVENTION: Small Peptides Having Potent Anti-Angiogenic Activity
3. FILE REFERENCE: 1781-170P
4. CURRENT APPLICATION NUMBER: US/09/385,442
5. CURRENT FILING DATE: 1999-08-30
6. EARLIER APPLICATION NUMBER: 60/099,313
7. EARLIER FILING DATE: 1999-09-04
8. NUMBER OF SEQ ID NOS: 50
9. SOFTWARE: PatentIn Ver. 2.0
10. SEQ ID NO 32
11. LENGTH: 16
12. TYPE: PRT
13. ORGANISM: mammalian
14. FEATURE:
15. OTHER INFORMATION: Endo-4
16. US-09-385-442-32
17.
18. Query Match          27.2%; Score 94; DB 3; Length 16;
19. Best Local Similarity 87.5%; Pred. No. 4e+06;
20. Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
21.
22. QY      48 HPTWPKRSVHWGSDPN 63
23.         ||| ||||| ||||| :
24. Db       1 HPAPWPKRSVHWGSDPS 16

```

Search completed: September 10, 2003, 14:13:05
Job time : 22 secs

```
QY      48 HPTWPKSVWHGSDPN 63
          |||||:
Db      1 HPWPQKSVWHGSDPS 16
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

```
Run on:      September 10, 2003, 14:09:48 ; Search time 41 Seconds
              (without alignments)
              243.897 Million cell updates/sec
```

Title: US-09-822-540A-1

Sequence: 1 ADRAVPVIVNLKDELLFPW.....DVLRHPTWQKSVWHGSDPN 63

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

Database :

A_Geneseq_19um303.*

1:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24:	/SDSL1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	346	100.0	63	23	ABB08121	Human endostatin F
2	346	100.0	178	21	AAV94324	Alternate human en
3	346	100.0	178	22	AAU00301	Human Endostatin(T
4	346	100.0	179	22	AAU00301	Human Endostatin(T
5	346	100.0	180	22	AAU00899	Human Endostatin(T
6	346	100.0	181	22	AAU00898	Human Endostatin(T
7	346	100.0	182	21	AA28399	Human endostatin.
8	346	100.0	182	21	AAV94323	Human endostatin F
9	346	100.0	182	21	AAV59622	Human endostatin F

10	346	100.0	182	22	AAU00897	Human Endostatin(T
11	346	100.0	182	22	AAU77951	Amino acid sequenc
12	346	100.0	183	20	AAV08693	Human endostatin p
13	346	100.0	183	20	AAV02113	SEQ ID 76 of W0991
14	346	100.0	183	21	AAAB30493	Amino acid sequenc
15	346	100.0	183	21	AAAB16451	Human endostatin p
16	346	100.0	183	21	AAV90771	Human angiogenesis
17	346	100.0	183	21	AAV70252	Human angiogenesis
18	346	100.0	183	22	AAU00896	Human endostatin(T
19	346	100.0	183	22	AAAB9379	Human endostatin S
20	346	100.0	183	23	AAAB79901	Human endostatin p
21	346	100.0	183	23	AAU97112	Human endostatin
22	346	100.0	183	23	AAAM49503	Human endostatin p
23	346	100.0	183	23	AAAM48895	Human endostatin p
24	346	100.0	183	24	AAAG79753	Human endostatin..
25	346	100.0	184	22	AAAG78717	Human vascular end
26	346	100.0	184	23	ABG31794	Murine endostatin
27	346	100.0	193	21	AAAM90877	Human HMW endostat
28	346	100.0	193	23	AAU76690	Synthetic plaamid
29	346	100.0	195	21	AAAM00874	Human HMW endostat
30	346	100.0	216	21	AAAB30495	Amino acid sequenc
31	346	100.0	275	23	AAU76699	Synthetic plaamid
32	346	100.0	310	23	AAU76698	Human collagen XV
33	346	100.0	513	23	ABG73586	Human endostatin/T
34	346	100.0	682	23	ABP41878	Human ovarian anti
35	346	100.0	684	18	AAAM26377	Human alpha-1 colla
36	346	100.0	684	20	AAV25113	Human alpha1 (XVII
37	346	100.0	684	23	AAU77357	Human collagen tYP
38	346	100.0	1301	20	AAAM2236	Human alpha-1 (XVI
39	346	100.0	1336	20	AAV08694	Human collagen 18
40	346	100.0	1336	24	ABP96308	Human endostatin p
41	346	100.0	1516	23	ABP68617	Human pancreatic c
42	346	100.0	1516	23	ABBB3471	Human collagen XVI
43	342	98.8	183	22	AAAB49810	Human endostatin p
44	335	96.8	271	21	AAAB0407	A human angiogenes
45	302	87.3	184	21	AAV70255	Canine angiogenesis

ALIGNMENTS

RESULT 1
ABB08121
ID ABB08121 standard; protein; 63 AA

DT 10-SEP-2002 (first entry)

DE Human endostatin polypeptide fragment JCK-362 (residues 52-114)

KM Endostatin; endothelial; cell proliferation; angiogenesis; cytostatic
KM antiarthritic; antiarteriosclerotic; ophthalmological; antidiabetic;
KM antitumor; vulnery; gynaecological; human.

Homo sapiens.

FH	Key	Location/Qualifiers
23	23	23

FT	/note=	"JJC-367 peptide (ABB08122), specifically
FT		claimed fragment"

/note= "C-terminal amide"

PN US2002058620-A1.

PD 16-MAY-2002

30-MAR-2001; 2001US-0822540.

PR 03-APR-2000; 2000US-194561P.

PA (PHOE-) PHOENIX PHARMACOLOGIES INC.

```

XX
PI Chang J;
XX
DR MPI; 2002-499519/53.
XX
PT New purified polypeptide useful for reducing tumor volume in an animal
XX PT complies a specified amino acid sequence
XX
PS Claim 1; Fig 1; 19pp; English.
XX
CC The invention relates to endostatin polypeptide fragments that can be
CC used as inhibitors of endothelial cell proliferation and angiogenesis.
CC The polypeptide inhibitors are useful for reducing tumor volume in an
CC animal; for inhibiting the proliferation of endothelial cells in an
CC animal; for targeting and binding purified polypeptide to an angiogenesis
CC receptor; inhibiting metastasis, for identifying, isolating or purifying
CC receptor proteins. They are also useful in therapy for treating diseases
CC such as atherosclerosis, haemangioma, leukemia, myocardial angiogenesis,
CC neovascular glaucoma, diabetic retinopathy, retrolental fibroplasia,
CC arthritis, diabetic neovascularisation, macular degeneration, wound
CC healing, peptic ulcer, fracture, keloid, hematopoiesis, ovulation,
CC menestration and placentaion in human; and for producing monoclonal and
CC polyclonal antibodies. The present sequence represents a human endostatin
CC polypeptide fragment, JCR-362.
XX
SQ Sequence 63 AA;
XX
Query Match 100.0%; Score 346; DB 23; Length 63;
Best Local Similarity 100.0%; Pred. No. 3, 2e-39;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ADRAAVPIVNLKDELLFPGSEGLPKRGARIFSGDKDVLRRHPTWQKSWHGS 60
DB 1 ADRAAVPIVNLKDELLFPGSEGLPKRGARIFSGDKDVLRRHPTWQKSWHGS 60
XX
QY 61 DPN 63
DB 61 DPN 63
XX
RESULT 2
AA94324
ID AAY94324 standard; Protein; 178 AA.
XX
AC AAY94324;
XX
DE 11-AUG-2000 (first entry)
XX
DE Alternate human endostatin protein.
XX
KW Human; endothelial cell proliferation inhibitor; collagen XVIII;
KW angiogenesis inhibitor; anti-tumour; cytostatic; antiproliferative;
KW vasotrophic; dermatological; ophthalmological; vulnery;
KW antidiabetic; antidiabetic; haemostatic; contraceptive;
KW ocular angiogenesis; atherosclerosis; scleroderma;
KW myocardial angiogenesis; telangiectasia; angiodioma;
KW wound granulation.
XX
OS Homo sapiens.
XX
PN WO200026368-A2.
XX
PD 11-MAY-2000.
XX
PP 01-NOV-1999; 99WO-US25605.
XX
PR 30-OCT-1998; 98US-0106343.
XX PR 20-MAY-1999; 99US-0315689.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI O'Reilly MS, Folkman MJ;
XX
XX

```

```

DR MPI; 2000-365617/31.
XX
DR N-PSDB; AAA27005.
XX
PT Novel endostatin capable of inhibiting endothelial cell proliferation
XX PT and angiogenesis, useful for treating angiogenesis-dependent cancers
XX PT and as birth control agents
XX
PS Claim 3; Page 39; 68pp; English.
XX
CC The present sequence is an alternate functional endostatin
CC protein. When the human endostatin gene sequence AAA27004 is
CC recombinantly expressed, an observable doublet of protein results, both
CC versions of which are functional endostatin proteins. The present
CC endostatin variant is the same as the protein encoded by AAA27004 minus
CC the first four amino acids. Recombinant mouse endostatin (20 mg/kg) was
CC administered subcutaneously to mice implanted with Lewis lung carcinomas.
CC There was tumor mass regression non-detectable levels after 12 days of
CC therapy due to the angiogenesis inhibitory activity of endostatin. Thus
CC the protein is useful for treatment of angiogenesis-dependent cancers.
CC The polynucleotide and polypeptide sequences of this endostatin are
CC useful for treating and diagnosis of tumours, ocular angiogenic
CC diseases, Osler-Webber syndrome, myocardial angiogenesis, plaque
CC neovascularisation, telangiectasia, haemophilic joints, angiodioma
CC and wound granulation, for treatment of diseases related to excessive or
CC abnormal stimulation of endothelial cells e.g. intestinal adhesions,
CC atherosclerosis, scleroderma. The protein may also be useful as a birth
CC control agent by reducing or preventing uterine vascularisation. The
CC gene for endostatin may be isolated from cells or tissue that express
CC high levels of endostatin, eg. tumour cells, by generating cDNA from
CC mRNA using reverse transcriptase and then amplifying the DNA sequence.
XX
SQ Sequence 178 AA;
XX
Query Match 100.0%; Score 346; DB 21; Length 178;
Best Local Similarity 100.0%; Pred. No. 1, 2e-38;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ADRAAVPIVNLKDELLFPGSEGLPKRGARIFSGDKDVLRRHPTWQKSWHGS 60
DB 60 ADRAAVPIVNLKDELLFPGSEGLPKRGARIFSGDKDVLRRHPTWQKSWHGS 119
XX
QY 61 DPN 63
DB 120 DPN 122
XX
RESULT 3
AAU00900
ID AAU00900 standard; Protein; 178 AA.
XX
AC AAU00900;
XX
DE 04-JUL-2001 (first entry)
XX
DE Human Endostatin(TM) N-terminal deletion mutant protein#2.
XX
KW Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;
KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubecosis;
KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
KW retinopathy of prematurity; macular corneal graft rejection;
KW neovascular glaucoma; retrolental fibroplasia; Osler-Webber Syndrome;
KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
KW haemophilic joint; angiodioma; wound granulation; variant;
KW mutant; mutcin.
XX
OS Homo sapiens.
XX
PN WO200119989-A2.
XX
PD 22-MAR-2001.
XX
XX

```

PF 14-SEP-2000; 2000WO-US25166.
 XX
 PR 14-SEP-1999; 99US-0153698.
 XX
 PA (ENTR-) ENTREMED INC.
 XX
 PI Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;
 PI Bermejo LL, Mistry FR, Shepard SR, Schrimsher JL;
 XX
 XX WPI: 2001-244802/25.
 DR N-PSDB; AAS00868.
 XX
 DR
 XX Producing Endostatin protein for treating angiogenesis mediated
 PT diseases such as solid tumours, comprises recombinantly producing the
 PT protein using an expression system, and recovering and purifying the
 PT protein -
 PS Claim 5; Page 33; 67pp; English.
 XX
 XX The sequence represents Human Endostatin(TM) N-terminal deletion mutant
 CC protein lacking the N-terminal 4 amino acids and the C-terminal lysine,
 CC a natural variant recovered from fermentations of *Pichia pastoris*
 CC cultures harbouring a expression plasmid containing the Endostatin(TM)
 CC DNA sequence given in AAS00868. The new method of the invention is
 CC useful for producing, recovering and purifying Endostatin(TM) from
 CC biological sources, such as biological fluids, tissues, cells, culture
 CC media, and fermentation media. Endostatin(TM) is useful for treating
 CC angiogenesis mediated diseases such as solid tumours, blood borne
 CC tumours, leukaemias, tumour metastases, benign tumours, e.g. haemangioma,
 CC acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas,
 CC rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
 CC e.g., diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, colon
 CC cancer, retrolental fibroplasia, rubecosis, Osler-Webber Syndrome,
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,
 CC haemophilic joints, angiodiroma, and wound granulation. Endostatin(TM)
 CC is also useful for treating disease of excessive or abnormal stimulation
 CC of endothelial cells such as intestinal adhesions, atherosclerosis,
 CC scleroderma and hypertrophic scars. Higher yields of more purified, and
 CC biologically active Endostatin(TM) are obtained by the new method.
 CC Endostatin(TM) can be stored in buffers for extended periods of time, and
 CC also subjected to lyophilisation, while preserving biological activity.
 CC Centrifugation of broth from fermentation steps in production is avoided,
 CC preventing unwanted potential cellular lysis and contamination with
 CC additional proteins, pigments, enzymes and other cellular chemicals and
 CC debris.
 CC
 XX Sequence 178 AA;
 SQ
 Query Match 100.0%; Score 346; DB 22; Length 178;
 Best Local Similarity 100.0%; Pred. No. 1.2e-38;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ADRAAVPIVNLKDELFPSEWALFSGSGPLKPGARIFSPDGKDVLRHPTWPOKXVWMS 60
 DB 60 ADRAAVPIVNLKDELFPSEWALFSGSGPLKPGARIFSPDGKDVLRHPTWPOKXVWMS 119
 OY 61 DPN 63
 DB 120 DPN 122
 XX
 XX RESULT 4
 XX AAU00901
 XX ID AAU00901 standard; Protein; 179 AA.
 XX AC AAU00901;
 XX XX
 XX DT 04-JUL-2001 (first entry)
 XX DE Human Endostatin(TM) N-terminal mutant protein#1.
 XX KW Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;

KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;
 KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubecosis;
 KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
 KW retinopathy of prematurity; macular corneal graft rejection;
 KW neovascular glaucoma; retrolental fibroplasia; Osler-Webber Syndrome;
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 KW haemophilic joint; angiodiroma; wound granulation; mutant; mutein.
 XX
 OS Homo sapiens.
 XX
 XX WO200119989-A2.
 XX
 XX 22-MAR-2001.
 XX PD
 XX 14-SEP-2000; 2000WO-US25166.
 XX PF
 XX 14-SEP-1999; 99US-0153698.
 XX
 XX
 XX (ENTR-) ENTREMED INC.
 XX
 XX Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;
 PI Bermejo LL, Mistry FR, Shepard SR, Schrimsher JL;
 XX
 XX WPI: 2001-244802/25.
 DR N-PSDB; AAS00868.
 XX
 DR
 XX Producing Endostatin protein for treating angiogenesis mediated
 PT diseases such as solid tumours, comprises recombinantly producing the
 PT protein using an expression system, and recovering and purifying the
 PT protein -
 PS Claim 5; Page 32; 67pp; English.
 XX
 XX The sequence represents a Human Endostatin(TM) N-terminal deletion
 CC mutant lacking the N-terminal 4 amino acids. The new method of the
 CC invention is useful for producing, recovering and purifying Endostatin
 CC (TM) from biological sources, such as biological fluids, tissues, cells,
 CC culture media, and fermentation media. Endostatin(TM) is useful for
 CC treating angiogenesis mediated diseases such as solid tumours, blood
 CC borne tumours, leukaemias, tumour metastases, benign tumours, e.g.
 CC haemangioma, acoustic neuromas, neurofibromas, trachomas, and pyogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
 CC e.g., diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, colon
 CC cancer, retrolental fibroplasia, rubecosis, Osler-Webber Syndrome,
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,
 CC haemophilic joints, angiodiroma, and wound granulation. Endostatin(TM)
 CC is also useful for treating disease of excessive or abnormal stimulation
 CC of endothelial cells such as intestinal adhesions, atherosclerosis, and
 CC scleroderma and hypertrophic scars. Higher yields of more purified, and
 CC biologically active Endostatin(TM) are obtained by the new method.
 CC Endostatin(TM) can be stored in buffers for extended periods of time, and
 CC also subjected to lyophilisation, while preserving biological activity.
 CC Centrifugation of broth from fermentation steps in production is avoided,
 CC preventing unwanted potential cellular lysis and contamination with
 CC additional proteins, pigments, enzymes and other cellular chemicals and
 CC debris.
 CC
 XX Sequence 179 AA;
 SQ
 Query Match 100.0%; Score 346; DB 22; Length 179;
 Best Local Similarity 100.0%; Pred. No. 1.2e-38;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ADRAAVPIVNLKDELFPSEWALFSGSGPLKPGARIFSPDGKDVLRHPTWPOKXVWMS 60
 DB 60 ADRAAVPIVNLKDELFPSEWALFSGSGPLKPGARIFSPDGKDVLRHPTWPOKXVWMS 119
 OY 61 DPN 63
 DB 120 DPN 122
 XX

RESULT 5
AAU00899
ID AAU00899 standard; Protein; 180 AA.
XX
AC AAU00899;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human Endostatin(TM) C-terminus minus 3 protein.
XX
KW Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;
KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubecosis;
KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
KW retinopathy of prematurity; macular corneal graft rejection;
KW neovascular glaucoma; retrolental fibroplasia; Osler-Weber Syndrome;
KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
KW haemophilic joint; angiofibroma; wound granulation; variant;
KW C-terminus minus 3 protein.
XX
OS Homo sapiens.
XX
PN WO200119989-A2.
XX
PD 22-MAR-2001.
XX
PF 14-SEP-2000; 2000WO-US25166.
XX
PR 14-SEP-1999; 99US-0153698.
XX
PA (ENTR-) ENTREMED INC.
XX
PI Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;
PI Bermejo LL, Mistry FR, Shepard SR, Schrimsher JL;
XX N-PSDB; AAS00867.
XX
DR WPI; 2001-244802/25.
XX
PT Producing Endostatin protein for treating angiogenesis mediated
PT diseases such as solid tumours, comprises recombinantly producing the
PT protein using an expression system, and recovering and purifying the
PT protein -
XX
PS Claim 5; Page 30; 67pp; English.
XX
CC The sequence represents Human Endostatin(TM) C-terminus minus 3
CC protein, a natural variant lacking the C-terminal 3 amino acids of
CC Endostatin(TM) recovered from fermentations of Pichia
CC pastoria cultures harbouring a expression plasmid containing the
CC Endostatin(TM) DNA sequence given in AAS00867. The new method of the
CC invention is useful for producing, recovering and purifying Endostatin
CC (TM) from biological sources, such as biological fluids, tissues, cells,
CC culture media, and fermentation media. Endostatin(TM) is useful for
CC treating angiogenesis mediated diseases such as solid tumours, blood
CC borne tumours, leukaemias, tumour metastases, benign tumours, e.g.
CC haemangioma, acoustic neuroma, neurofibromas, trachomas, and pyogenic
CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
CC e.g., diabetic retinopathy, retinopathy of prematurity, macular
CC degeneration, corneal graft rejection, neovascular glaucoma, colon
CC cancer, retrolental fibroplasia, rubecosis, Osler-Webber Syndrome,
CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,
CC is also useful for treating disease of excessive or abnormal stimulation
CC of endothelial cells such as intestinal adhesions, atherosclerosis,
CC scleroderma and hypertrophic scars. Higher yields of more purified, and
CC biologically active Endostatin(TM) are obtained by the new method.
CC Endostatin(TM) can be stored in buffers for extended periods of time, and
CC also subjected to lyophilisation, while preserving biological activity.
CC Centrifugation of broth from fermentation steps in production is avoided,
CC preventing unwanted potential cellular lysis and contamination with
CC additional proteins, pigments, enzymes and other cellular chemicals and

CC debris.
XX
SQ Sequence 180 AA;
XX
Query Match 100.0%; Score 346; DB 22; Length 180;
Best Local Similarity 100.0%; Pred. No. 1-2e-39;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ADRAVPYVNLKDELLFSGSEGLKPGARIFSGDKDYLRRPTPQKSWHGS 60
DB 64 ADRAVPYVNLKDELLFSGSEGLKPGARIFSGDKDYLRRPTPQKSWHGS 123
QY 61 DPN 63
DB 124 DPN 126
XX
RESULT 6
AAU00898
ID AAU00898 standard; Protein; 181 AA.
XX
AC AAU00898;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human Endostatin(TM) C-terminus minus 2 protein.
XX
KW Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;
KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubecosis;
KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
KW retinopathy of prematurity; macular corneal graft rejection;
KW neovascular glaucoma; retrolental fibroplasia; Osler-Weber Syndrome;
KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
KW haemophilic joint; angiofibroma; wound granulation; variant;
KW C-terminus minus 2 protein.
XX
OS Homo sapiens.
XX
PN WO200119989-A2.
XX
PD 22-MAR-2001.
XX
PF 14-SEP-2000; 2000WO-US25166.
XX
PR 14-SEP-1999; 99US-0153698.
XX
PA (ENTR-) ENTREMED INC.
XX
PI Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;
PI Bermejo LL, Mistry FR, Shepard SR, Schrimsher JL;
XX N-PSDB; AAS00867.
XX
DR WPI; 2001-244802/25.
XX
PT Producing Endostatin protein for treating angiogenesis mediated
PT diseases such as solid tumours, comprises recombinantly producing the
PT protein using an expression system, and recovering and purifying the
PT protein -
XX
PS Claim 5; Page 30; 67pp; English.
XX
CC The sequence represents Human Endostatin(TM) C-terminus minus 2
CC protein, a natural variant lacking the C-terminal 2 amino acids of
CC Endostatin(TM) recovered from fermentations of Pichia
CC pastoria cultures harbouring a expression plasmid containing the
CC Endostatin(TM) DNA sequence given in AAS00867. The new method of the
CC invention is useful for producing, recovering and purifying Endostatin
CC (TM) from biological sources, such as biological fluids, tissues, cells,
CC culture media, and fermentation media. Endostatin(TM) is useful for
CC treating angiogenesis mediated diseases such as solid tumours, blood
CC borne tumours, leukaemias, tumour metastases, benign tumours, e.g.

CC haemangioma, acoustic neuromas, neurofibromas, trachomas, and pyogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
 CC e.g., diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, colon
 CC cancer, retrorenal fibroplasia, rubecosis, Osler-Weber Syndrome,
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,
 CC haemophilic joints, angiodioma, and wound granulation. Endostatin(TM)
 CC is also useful for treating disease of excessive or abnormal stimulation
 CC of endothelial cells such as intestinal adhesions, atherosclerosis,
 CC scleroderma and hypertrophic scars. Higher yields of more purified, and
 CC biologically active Endostatin(TM) are obtained by the new method.
 CC Endostatin(TM) can be stored in buffers for extended periods of time, and
 CC also subjected to lyophilisation, while preserving biological activity.
 CC Centrifugation of broth from fermentation steps in production is avoided,
 CC preventing unwanted potential cellular lysis and contamination with
 CC additional proteins, pigments, enzymes and other cellular chemicals and
 CC debris.

CC Sequence 181 AA;

Query Match 100.0%; Score 346; DB 22; Length 181;

Best Local Similarity 100.0%; Pred. No. 1.2e-38;

Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVMHGS 60

DB 64 ADRAAVPIVNLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVMHGS 123

QY 61 DPN 63

DB 124 DPN 126

RESULT 7

AAZ8399

ID AAB28399 standard; Protein; 182 AA.

AC AAB28399;

XX 19-FEB-2001 (first entry)

DE Human endostatin.

XX Human; endostatin; cytosolic; antiproliferative;

KM vascular endothelial growth factor; VEGF; antibody; VEGF2 receptor;

XX cancer; vascularised solid tumour.

OS Homo sapiens.

XX WO200064946-A2.

XX 02-NOV-2000.

XX 28-APR-2000; 2000WO-US11367.

XX 28-APR-1999; 99US-0131432.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Thorpe PE, Brecken RA;

XX WPI; 2000-687317/67.

XX Immunogenic composition for the treatment and diagnosis of cancer

XX comprises an anti-VEGF (vascular endothelial growth factor) antibody

XX binding the same epitope as the monoclonal antibody ATCC PTA 1595 -

XX Example 10; Page 291-292; 298pp; English.

XX The present invention relates to anti-Vascular Endothelial Growth Factor

XX (VEGF) antibodies that bind to the same epitope as the monoclonal

XX antibody ATCC PTA 1595 and which significantly inhibit VEGF binding to

XX the VEGF receptor VEGFR2, without inhibiting VEGF binding to the VEGF

CC receptor VEGFR1. The present sequence is human endostatin. Endostatin

CC may be conjugated onto the anti-VEGF antibodies of the present invention.

CC The anti-VEGF antibodies of the present invention are useful for the

CC treatment and diagnosis of cancer, especially vascularised solid tumours.

CC Sequence 182 AA;

Query Match 100.0%; Score 346; DB 21; Length 182;

Best Local Similarity 100.0%; Pred. No. 1.2e-38;

Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADRAAVPIVNLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVMHGS 60

DB 64 ADRAAVPIVNLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVMHGS 123

QY 61 DPN 63

DB 124 DPN 126

RESULT 8

AAZ94323

ID AAY94323 standard; Protein; 182 AA.

AC AAY94323;

XX 11-AUG-2000 (first entry)

DE Human endostatin protein.

XX Human; endothelial cell proliferation inhibitor; collagen XVII;

KM angiogenesis inhibitor; anti-tumour; cytosolic; antiproliferative;

KM vasotrophic; dermatological; ophthalmological; vulvovaginal;

KM antiarteriosclerotic; antidiabetic; haemostatic; contraceptive;

KM ocular angiogenic disease; atherosclerosis; scleroderma;

KM myocardial angiogenesis; telangiectasia; angiodioma;

KM wound granulation.

OS Homo sapiens.

XX WO200026368-A2.

XX 11-MAY-2000.

XX 01-NOV-1999; 99WO-US25605.

XX 30-OCT-1998; 98US-0106343.

XX 20-MAY-1999; 99US-0315689.

XX (CHIL-) CHILDRENS MEDICAL CENT.

XX O'Reilly MS, Folkman MJ;

XX WPI; 2000-365617/31.

XX N-PSDB; AAA27004.

XX Novel endostatin capable of inhibiting endothelial cell proliferation

XX and as birth control agents

XX Claim 2; Page 38; 68pp; English.

XX The present sequence is an endostatin protein which is the carboxy

XX terminal protein of human collagen XVII. Recombinant mouse endostatin

XX (20 mg/kg) was administered subcutaneously to mice implanted with Lewis

XX lung carcinomas. There was tumour mass regression non-detectable levels

XX after 12 days of therapy due to the angiogenesis inhibitory activity of

XX endostatin. Thus the protein is useful for treatment of angiogenesis-

XX dependent cancers. The polynucleotide and polypeptide sequences of this

XX endostatin are useful for treating and diagnosis of tumours, ocular

XX angiogenic diseases, Osler-Weber syndrome, myocardial angiogenesis,

XX plaque neovascularisation, telangiectasia, haemophilic joints,

XX angiodioma and wound granulation, for treatment of diseases related to

CC excessive or abnormal stimulation of endothelial cells e.g. intestinal
 CC adhesions, atherosclerosis, scleroderma. The protein may also be useful
 CC as a birth control agent by reducing or preventing uterine
 CC vascularization. The gene for endostatin may be isolated from cells or
 CC tissue that express high levels of endostatin, eg. tumour cells, by
 CC generating cDNA from mRNA using reverse transcriptase and then amplifying
 CC the DNA sequence.
 CC
 SO Sequence 182 AA;
 Query Match 100.0%; Score 346; DB 21; Length 182;
 Best Local Similarity 100.0%; Pred. No. 1.2e-38;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ADRAAIVIVNLKDELFPSEWALFSGSGPLKPGARIFSPDGKDVLRHPTWPKSVWHS 60
 DB 64 ADRAAIVIVNLKDELFPSEWALFSGSGPLKPGARIFSPDGKDVLRHPTWPKSVWHS 123
 QY 61 DPN 63
 DB 124 DPN 126
 RESULT 9
 ID AAY59622 standard; protein; 182 AA.
 AC AAY59622;
 XX
 DT 14-MAR-2000 (first entry)
 DE Human endostatin protein fragment.
 XX
 KM Endostatin; scatter factor activity; human; tubulogenesis; psoriasis;
 KM metastatic cancer; tumorigenesis; ocular angiogenic disease;
 KM rheumatoid arthritis; Osler-Weber syndrome; telangiectasia;
 KM haemophilic joint; angiodioma; wound granulation.
 XX
 OS Homo sapiens.
 PN MO962944-A2.
 PD 09-DEC-1999.
 PF 03-JUN-1999; 99MO-US12278.
 PR 03-JUN-1998; 98US-0087890.
 PR 10-JUL-1998; 98US-0092393.
 PR 01-SEP-1998; 98US-0098790.
 PA (CHIL-) CHILDRENS MEDICAL CENT.
 PI Javaherian K, Folkman MJ;
 DR WPI; 2000-072833/06.
 PT New endostatin oligomers, used for treating e.g. tumours -
 PS Disclosure; Page 6; 44pp; English.
 CC This sequence is a fragment of the human endostatin protein. Endostatin
 CC is an approximately 20kD C-terminal globular domain of the collagen-like
 CC protein collagen XVIII. Protein oligomers consisting of more than one
 CC endostatin monomer have anti-tubulogenic effects and induce
 CC reorganization of the actin cytoskeleton. The oligomer has scatter factor
 CC activity. The oligomers induce the destruction of tubular lumens and
 CC elongation of cells, and inhibit tubulogenesis and tumorigenesis. The
 CC oligomers can also be used to treat metastatic cancers, tumours,
 CC rheumatoid arthritis, psoriasis, ocular angiogenic disease, Osler-Weber
 CC syndrome, plaque neovascularisation, telangiectasia, haemophilic
 CC joints, angiodioma and wound granulation. The oligomers can also be
 CC used to treat diseases that have angiogenesis as a pathological
 CC consequence e.g. ulcers. The endostatin oligomers can also be used to

CC develop affinity columns for isolating antibodies or receptors. Passive
 CC antibody therapy using antibodies that specifically bind endostatin
 CC oligomers can be used to modulate morphogenic processes such as
 CC metastatic cancer as well as angiogenesis-dependent processes such as
 CC reproduction, development, wound healing, tissue repair, and
 CC angiogenesis-dependent diseases. Also, antisera directed to the Fab
 CC regions of endostatin oligomer antibodies can be administered to block
 CC the ability of endogenous endostatin oligomer antisera to bind endostatin
 CC oligomers.
 CC
 SO Sequence 182 AA;
 Query Match 100.0%; Score 346; DB 21; Length 182;
 Best Local Similarity 100.0%; Pred. No. 1.2e-38;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ADRAAIVIVNLKDELFPSEWALFSGSGPLKPGARIFSPDGKDVLRHPTWPKSVWHS 60
 DB 64 ADRAAIVIVNLKDELFPSEWALFSGSGPLKPGARIFSPDGKDVLRHPTWPKSVWHS 123
 QY 61 DPN 63
 DB 124 DPN 126
 RESULT 10
 ID AAU00897 standard; protein; 182 AA.
 AC AAU00897;
 XX
 DT 04-JUL-2001 (first entry)
 DE Human Endostatin(TM) C-terminus minus 1 protein.
 XX
 KM Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
 KM blood borne tumour; leukaemia; tumour metastasis; benign tumour;
 KM haemangioma; acoustic neuroma; neurofibroma; trachoma; tubecosis;
 KM pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
 KM ocular angiogenic disease; diabetic retinopathy; macular degeneration;
 KM retinopathy of prematurity; macular corneal graft rejection;
 KM neovascular glaucoma; retrolental fibroplasia; Osler-Weber Syndrome;
 KM myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 KM haemophilic joint; angiodioma; wound granulation; variant;
 KM C-terminus minus 1 protein.
 XX
 OS Homo sapiens.
 PN WO200119989-A2.
 PD 22-MAR-2001.
 PF 14-SEP-2000; 2000WO-US25166.
 PR 14-SEP-1999; 99US-0153698.
 PA (ENTR-) ENTREMED INC.
 PI Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;
 DR Bermejo LL, Mistry FR, Shepard SR, Schrimsher JL;
 DR WPI; 2001-244802/25.
 PT N-P5DB; AAS00897.
 CC Producing Endostatin protein for treating angiogenesis mediated
 CC diseases such as solid tumours, comprises recombinantly producing the
 CC protein using an expression system, and recovering and purifying the
 CC protein -
 PS Claim 5; Page 30; 67pp; English.
 CC The sequence represents Human Endostatin(TM) C-terminus minus 1
 CC protein, a natural variant lacking the C-terminal amino acid of

CC Endostatin(TM) recovered from fermentations of Pichia
 CC pastoris cultures harbouring a expression plasmid containing the
 CC Endostatin(TM) DNA sequence given in AAS00867. The new method of the
 CC invention is useful for producing, recovering and purifying Endostatin
 CC (TM) from biological sources, such as biological fluids, tissues, cells,
 CC culture media, and fermentation media. Endostatin(TM) is useful for
 CC treating angiogenesis mediated diseases such as solid tumours, blood
 CC borne tumours, leukemias, tumour metastases, benign tumours, e.g.
 CC haemangioma, acoustical neuromas, neurofibromas, trachomas, and pyogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
 CC e.g., diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, colon
 CC cancer, retrolental fibroplasia, rubeosis, Osler-Weber Syndrome,
 CC mycobacterial angiogenesis, plaque neovascularisation, telangiectasia,
 CC haemophilic joints, angiodioma, and wound granulation. Endostatin(TM)
 CC is also useful for treating disease of excessive or abnormal stimulation
 CC of endothelial cells such as intestinal adhesions, atherosclerosis,
 CC scleroderma and hypertrophic scars. Higher yields of more purified, and
 CC biologically active Endostatin(TM) are obtained by the new method.
 CC Endostatin(TM) can be stored in buffers for extended periods of time, and
 CC also subjected to lyophilisation, while preserving biological activity.
 CC Centrifugation of broth from fermentation steps in production is avoided,
 CC preventing unwanted potential cellular lysis and contamination with
 CC additional proteins, pigments, enzymes and other cellular chemicals and
 CC debris.

CC Sequence 182 AA:

Query Match 100.0%; Score 346; DB 22; Length 182;
 Best Local Similarity 100.0%; Pred. No. 1.2e-38;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADRAAVPIVNLKDELPFSWEALFSGSGPLKPGARIFSPDGKDYLRHPTWPKSWMGS 60
 DB 64 ADRAAVPIVNLKDELPFSWEALFSGSGPLKPGARIFSPDGKDYLRHPTWPKSWMGS 123

OY 61 DPN 63
 DB 124 DPN 126

RESULT 11
 ID AAU77951 standard; Protein; 182 AA.

XX AC AAU77951;

DT 02-JUL-2002 (first entry)

DE Amino acid sequence for human endostatin.

KM Human; immunoglobulin; anti-vascular endothelial growth factor antibody;
 KM anti-VEGF antibody; monoclonal antibody 2C3 ATCC PTA 1595; VEGF receptor;
 KM VEGFR2; KDR/Flk-1; VEGFR1; Flt-1; angiogenesis; macular degeneration;
 KM ocular neovascular disease; cancer; vascularised solid tumour; AIDS;
 KM metastatic tumour; endothelial cell proliferation; inflammatory disorder;
 KM atherosclerosis; diabetic retinopathy; corneal graft rejection;
 KM acquired immune deficiency syndrome; infection; restenosis; fungal ulcer;
 KM sickle cell anaemia; endometriosis; endostatin.

OS Homo sapiens.

PN AU200179401-A.

PD 06-DEC-2001.

PF 12-OCT-2001; 2001AU-0079401.

PR 28-APR-2000; 2000AU-0048049.

XX 12-OCT-2001; 2001AU-0079401.

XX (TEXA) UNIV TEXAS SYSTEM.

PI Thorpe PE, Brecken RA;
 XX WPI; 2002-281368/33.

PT Immunoglobulin compositions for treating cancer by inhibiting
 PT angiogenesis and for delivering a diagnostic agent to tumour, comprises
 PT anti-vascular endothelial growth factor antibody attached to a
 PT biological agent

XX Example 10; Page 12-13 (Sequence listing); 300pp; English.

CC The present invention relates to antibody-based compositions comprising
 CC an immunoglobulin such as anti-vascular endothelial growth factor
 CC (VEGF) antibody (Ab) (or its antigen-binding fragment), attached to a
 CC biological agent, where the Ab binds to the same epitope as the
 CC monoclonal antibody (Mab) 2C3 ATCC PTA 1595, and significantly inhibits
 CC VEGF binding to the VEGF receptor VEGFR2 (KDR/Flk-1) without inhibiting
 CC VEGF binding to the VEGF receptor VEGFR1 (Flt-1). The compositions
 CC of the invention are useful in therapy, and diagnosis, for inhibiting
 CC angiogenesis in an animal having ocular neovascular disease or macular
 CC degeneration, and for delivering a biological agent to a vascularised
 CC tumour. The compositions can also be used for treating cancer and
 CC subjects at risk of developing a vascularised solid tumour, a metastatic
 CC tumour or metastases from a primary tumour. The composition is useful
 CC for specifically inhibiting VEGF-induced endothelial cell proliferation,
 CC without significantly inhibiting VEGF-induced macrophage, osteoclast or
 CC chondrocyte function. The compositions can be used for treating various
 CC diseases such as inflammatory disorders, atherosclerosis, diabetic
 CC retinopathy, restenosis, acquired immune deficiency syndrome (AIDS),
 CC blood borne tumours, corneal graft rejection, Crohn's disease, fungal
 CC ulcers, infections, sickle cell anaemia, and endometriosis. The present
 CC sequence represents human endostatin. Endostatin may be attached or
 CC functionally associated with anti-VEGF antibodies.

SO Sequence 182 AA:

Query Match 100.0%; Score 346; DB 23; Length 182;
 Best Local Similarity 100.0%; Pred. No. 1.2e-38;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADRAAVPIVNLKDELPFSWEALFSGSGPLKPGARIFSPDGKDYLRHPTWPKSWMGS 60
 DB 64 ADRAAVPIVNLKDELPFSWEALFSGSGPLKPGARIFSPDGKDYLRHPTWPKSWMGS 123

OY 61 DPN 63
 DB 124 DPN 126

RESULT 12
 ID AAY08693 standard; Protein; 183 AA.

XX AC AAY08693;

DT 10-AUG-1999 (first entry)

DE Human endostatin protein fragment.

KM Plasmogen; human; angiostatin; endostatin; gene therapy; vector;
 KM anti-angiogenic; attenuation; cytoskeletal; anti-diabetic; ophthalmology;
 KM tumour growth; solid tumour; diabetic retinopathy; retina.

OS Homo sapiens.

PN WO926480-A1.

PD 03-JUN-1999.

PF 20-NOV-1998; 98WO-US24950.

XX 20-NOV-1997; 97US-0975424.

PA (GENE-) GENETIX PHARM INC.
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
XX
PI Bachelot T, Leboulch P, Pawlinski RJ;
XX WPI, 1999-357696/30.
DR N-PSDB; AAX77719.
XX
XX
PT Anti-angiogenic gene therapy vectors
XX
PS Disclosure; Page 74-75; 83pp; English.
XX
CC This invention describes a novel viral gene therapy vector comprising a
CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen
CC from human or murine angiostatin, human or murine endostatin and
CC angiogenesis-inhibiting fusions and fragments, where the viral vector is
CC efficiently attenuated for use in human gene therapy. The products of
CC the invention have anti-angiogenic, cytostatic, anti-diabetic and
CC ophthalmological activity. The vector is used in gene therapy for
CC inhibiting tumour growth in humans harbouring a solid tumour. The vector
CC expresses an anti-angiogenic polypeptide. An additional use comprises
CC treatment of diabetic retinopathy, where the anti-angiogenic polypeptide
CC inhibits angiogenesis in the vicinity of the retina. The vector is
CC administered to cells ex vivo and then administered to the patient.
XX
XX
SQ Sequence 183 AA;
XX
XX
Query Match 100.0%; Score 346; DB 20; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.2e-38;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ADRAAVPIVNLKDELFPSEWALFSGSEGLPKGARIFSPDGKDVLRHPTWPKSVWHS 60
DB 64 ADRAAVPIVNLKDELFPSEWALFSGSEGLPKGARIFSPDGKDVLRHPTWPKSVWHS 123
XX
QY 61 DPN 63
DB 124 DPN 126
XX
XX
RESULT 13
AA02113
ID AAY02113 standard; Protein; 183 AA.
XX
AC AAY02113;
XX
XX 16-JUL-1999 (first entry)
XX
DE SEQ ID 76 of WO916889.
XX
KM Angiostatin; endostatin; interferon; thrombospondin;
KM anti-tumor-inducible protein; platelet factor 4; anti-angiogenic;
KM anti-tumor; multifunctional protein; angiogenic-mediated disease;
KM cancer; diabetic retinopathy; macular degeneration; arthritis;
KM tumor cell production.
XX
OS Homo sapiens.
XX
XX WO916889-A1.
XX
XX 08-APR-1999.
XX
XX 30-SEP-1998; 98WO-US20464.
XX
XX 01-OCT-1997; 97US-0060609.
XX
XX (SEAR) SEARLE & CO G D.
XX
XX Bolanowski MA, Caparon MH, Casperson GF, Gregory SA;
PI Klein BK, McKearn JP;
XX WPI, 1999-255098/21.
XX

PT New multifunctional proteins useful for treating angiogenic-mediated
PT diseases
XX
XX
XX Disclosure; Page 106-107; 121pp; English.
XX
XX
XX The specification describes multifunctional proteins which comprise
CC combinations of angiostatin, endostatin, interferon, thrombospondin,
CC interferon-inducible protein and platelet factor 4, and have
CC anti-angiogenic and/or anti-tumor activity. The multifunctional protein
CC may exhibit useful properties such as having similar or greater
CC biological activity when compared to a single factor or by having
CC improved half-life or decreased adverse side effects, or a combination
CC of these properties. The proteins can be used for treating an
CC angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
CC degeneration, or arthritis. They can also be used for inhibiting the
CC production of tumor cells (characteristic of lung, breast, ovarian,
CC prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,
CC hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
CC growth. The present sequence is used in the course of the invention.
XX
XX
SQ Sequence 183 AA;
XX
XX
Query Match 100.0%; Score 346; DB 20; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.2e-38;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ADRAAVPIVNLKDELFPSEWALFSGSEGLPKGARIFSPDGKDVLRHPTWPKSVWHS 60
DB 64 ADRAAVPIVNLKDELFPSEWALFSGSEGLPKGARIFSPDGKDVLRHPTWPKSVWHS 123
XX
QY 61 DPN 63
DB 124 DPN 126
XX
XX
RESULT 14
AAB30493
ID AAB30493 standard; Protein; 183 AA.
XX
AC AAB30493;
XX
XX 06-MAR-2001 (first entry)
XX
DE Amino acid sequence of human endostatin encoded by plasmid pMALCH15.
XX
XX Streptomyces sp. strain C5; SnpA; S. venezuelae; alpha-amylase;
XX endostatin; cancer; tumour growth; angiogenesis.
XX
OS Homo sapiens.
XX
XX WO200060945-A1.
XX
XX 19-OCT-2000.
XX
XX 12-APR-2000; 2000WO-US09747.
XX
XX 13-APR-1999; 99US-0129084.
XX
XX (MERI) MERCK & CO INC.
XX
XX Desautel CL, Strohl WR;
XX
XX WPI; 2000-686970/67.
XX
XX N-PSDB; AAC62023.
XX
XX Preparation of soluble recombinant endostatin involves transforming
PT Streptomyces host with expression vector comprising nucleotide
PT sequence encoding endostatin operably linked to linker and leader
PT peptide
XX
XX Example 1; Fig 6; 57pp; English.
XX
XX The present sequence represents human endostatin. The protein is
CC

CC expressed in Streptomyces. Leader sequences of Streptomyces sp. strain
 CC C5 Snpa and S. venezuelae alpha-amylase proteins are linked to the
 CC N-terminal of endostatin. This ensures that endostatin protein is
 CC produced as a secreted, soluble protein which needs no refolding, is
 CC stable in the fermentation broth and is produced in large quantities.
 CC The method is used for preparing soluble recombinant human, murine or
 CC primate endostatin, which is useful in the treatment of cancer,
 CC inhibition of tumour growth, inhibition of angiogenesis, isolation of
 CC receptors for endostatin and for identification of anti-angiogenic
 CC compounds in assays. The endostatin protein is produced as a secreted,
 CC soluble protein which needs no refolding, is stable in the fermentation
 CC broth and is produced in large quantities. Streptomyces are amenable
 CC for cultivation in large fermentations allowing for large quantities of
 CC soluble endostatin to be produced.

SQ Sequence 183 AA;
 Query Match 100.0%; Score 346; DB 21; Length 183;
 Best Local Similarity 100.0%; Pred. No. 1.2e-38;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADRAAVPIVNLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDYLRHPTWPKSVWHS 60
 DB 64 ADRAAVPIVNLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDYLRHPTWPKSVWHS 123

OY 61 DPN 63
 DB 124 DPN 126

RESULT 15
 AAB16451
 ID AAB16451 standard; Protein; 183 AA.
 XX
 AC AAB16451;
 DT 27-OCT-2000 (first entry)
 XX
 DE Human endostatin protein sequence.
 XX
 KW Angiogenesis-inhibiting protein receptor; angiogenesis; angiotatin;
 KW endostatin; plasminogen; laminin; treatment; wound healing; solid tumour;
 KW psoriasis; scleroderma; myocardial angiogenesis; Crohn's disease;
 KW cerebral collateral; arteriovenous malformation; rubecosis; cancer;
 KW diabetic retinopathy; arthritis; wound healing; peptic ulcer;
 KW Helicobacter related disease; fracture; cat scratch fever.
 XX
 OS Homo sapiens.
 XX
 PN WO200032631-A2.
 XX
 PD 08-JUN-2000.
 XX
 PF 06-DEC-1999; 99WO-US28897.
 XX
 PR 04-DEC-1998; 98US-0206059.
 XX
 PA (ENTR-) ENTREMED INC.
 XX
 PI MacDonald NJ, Sim KL;
 DR WPI; 2000-412290/35.
 XX
 PT New angiogenesis-inhibiting protein receptors, useful in methods for
 PT treating diseases and processes that are mediated by angiogenesis, such
 PT as solid tumours, psoriasis, scleroderma and myocardial angiogenesis -
 PS Disclosure; Figure 3; 100pp; English.
 XX
 CC This invention relates to angiogenesis-inhibiting protein receptors, and
 CC the DNA sequences encoding them. Angiogenesis is the generation of new
 CC blood vessels into a tissue, and normally occurs in wound healing,
 CC foetal and embryonal development and the formation of the corpus luteum.

CC endometrium and placenta. Angiotatin is a protein (see AAB16450 and
 CC AAA68202) involved in angiogenesis, and has an amino acid sequence
 CC similar to that of a plasminogen fragment (see murine plasminogen
 CC AAB16490). Angiotatin has the ability to inhibit angiogenesis.
 CC Endostatin is also an angiogenesis inhibiting protein (see AAB16451 and
 CC AAA68203). Sequences AAA68242 and AAB16522 represent coding and protein
 CC sequences of human laminin. Laminin is an angiotatin binding protein,
 CC and some of the peptides of the invention share homology with regions of
 CC laminin. Peptides AAB16452-816521 (excluding AAB16490) are the
 CC angiogenesis-inhibiting protein receptor fragments of the invention. The
 CC peptides bind either angiotatin or endostatin and can be used in methods
 CC for treating diseases and processes that are mediated by angiogenesis,
 CC such as solid tumours, psoriasis, scleroderma, myocardial angiogenesis,
 CC Crohn's disease, cerebral collateral, arteriovenous malformations,
 CC rubecosis, diabetic retinopathy, arthritis, wound healing, peptic ulcers,
 CC Helicobacter related diseases, fractures, placentalion and cat scratch
 CC fever. They are useful for the detection and prognosis of cancer. DNA
 CC sequences A628204-A628241 encode the peptides of the invention.

SQ Sequence 183 AA;

Query Match 100.0%; Score 346; DB 21; Length 183;

Best Local Similarity 100.0%; Pred. No. 1.2e-38;

Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADRAAVPIVNLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDYLRHPTWPKSVWHS 60
 DB 64 ADRAAVPIVNLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDYLRHPTWPKSVWHS 123

OY 61 DPN 63
 DB 124 DPN 126

Search completed: September 10, 2003, 14:12:30
 Job time : 42 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)